Australian Mental Health Outcomes and Classification Network

'Sharing Information to Improve Outcomes'

An Australian Government funded initiative

Health of the Nation Outcome Scales for Infants (HoNOSI) Field Trial

FINAL REPORT

2020

Acknowledgements

The Australia Mental Health Outcomes and Classification Network acknowledges that a number of people across the mental health sector have played an important role in the development and review of the Health of the Nation Outcomes Scales for Infants (HoNOSI). We would particularly like to acknowledge members of the Child and Adolescent Mental Health Information Development Expert Advisory Panel (CAMHIDEAP), the project working group and the clinicians and services who supported testing of the HoNOSI: Dr Peter Brann (Chair), Dr Titia Sprague, Dr Nick Kowalenko, Dr Simon Davies, Ms Margaret Hoyland, Dr Paul Denborough, Ms Margaret Jones, Ms Elida Meadows, Ms Helen Pepper, Dr Elisabeth Hoehn, Dr Anne Sved Williams, Dr Gordana Culjak, Dr Tim Coombs, Prof Philip Burgess, Ms Rosemary Dickson, Ms Lee Meredith, Ms Karen Hazell Raine, Ms Christina Hunt, Dr Karen Gaunson, Mr Ross Wood, Ms Lisa Phelan, Perinatal and Infant Mental Health/Child and Youth Mental Health Service, Northern Sydney LHD (NSW), Perinatal and Infant Mental Health Service, Western Sydney LHD (NSW), Western NSW LHD (NSW), Queensland Centre for Perinatal and Infant Mental Health - Children's Health Queensland Hospital and Health Service (QLD), Child and Youth Mental Health Service (CYMHS) - Children's Health Queensland Hospital and Health (QLD), Eastern Health - Child and Youth Mental Health Service (CYMHS) (VIC), Monash ELMHS (VIC), Alfred Health (CYMHS) (VIC), Helen Mayo House, Women and Children's Health Network (SA), Child and Adolescent Mental Health Services (CAMHS) (WA), Cairns Child and Youth Mental Health Service - Cairns and Hinterland Hospital and Health Service (QLD), Child and Youth Mental Health Service - Townsville Hospital and Health Service (QLD).

The development of the HoNOSI was funded by the Australian Government Department of Health.
All enquiries about this report should be directed to: Australian Mental Health Outcomes and Classification Network at email: contact@amhocn.org.

CONTENTS

EXECUTIVE SUMMARY

Background

Both the NOCC Strategic Review of routine outcomes in Australian mental health services and a subsequent targeted review identified a gap in measures addressing social, emotional and behavioural domains for pre-schoolers and infants. A working group from CAMHIDEAP built upon the existing structure of HoNOSCA to develop the Health of the Nation Outcome Scales for Infants (HoNOSI). The HoNOSI is intended as a routine outcome measure to support clinicians working with mental health of children aged 0-47 months. Face validity testing (Annex 1) showed HoNOSI was considered useful in measuring infant mental health outcomes. The current field trial was designed to further examine the reliability and validity of HoNOSI.

Method

Two studies were conducted; concurrent validity and inter-rater reliability, both involving a range of professionals from across Australia, who could be expected to use HoNOSI in routine clinical practice. Findings from these two studies were evaluated using the COnsensus based Standards for the selection of health status Measurement INstruments (COSMIN).

Results

The concurrent validity study examined the relationship between clinician's perceptions using HoNOSI and other measures on the infants and pre-schoolers currently in their treatment. The median age of the infants was 15 months with 108 cases included. Most clinicians had over five years' experience. The majority of scales used the full range of ratings. There was no missing data for any scale and no evidence of any ceiling effects. The majority of scales had *No Problem* as the most frequent rating, though the total score did not show any floor effect. The HoNOSI was statistically significantly correlated with the PIR-GAS, Clinical Worry and Severity Judgement ratings. The findings from the concurrent validity study were acceptable according to the COSMIN framework.

The inter-rater reliability of the HoNOSI was assessed by comparing independent clinical ratings on the same case vignettes. An initial pilot study of ratings completed by 81 clinicians on the same three vignettes showed that overall, estimates of inter-rater reliability were not acceptable in terms of COSMIN criteria. Using the consistency inter-rater reliability model, only two scales were *almost perfect*, one was *substantial*, five were *moderate*, six were *fair* and two were *slight*.

There are several possible explanations for the low inter-rater reliability estimates found during the pilot study. Examination of the vignettes revealed that they were all highly severe with the lowest median total score being almost twice as severe as the median total score of the real cases reported in the concurrent validity study. Similarly, the least severe vignette rating was above the 90th percentile for the real cases. Inter-rater reliability estimates are difficult to determine when there is little variability in the targets to be rated. While the vignette development checked for variation within scales, it did not address the resulting issue that there was substantial similarity in the overall ratings. Support for the hypothesis that a methodological problem contributed was the finding that

those scales with the greatest variability between vignettes tended to have the highest inter-rater reliability. Conversely, those scales with little variation between vignettes had lower inter-rater reliability estimates.

Subsequently, a HoNOSI Inter-rater reliability study was conducted with 45 raters completing five vignettes; three revised from the pilot study, with the addition of two newly created vignettes. Quadratic weighted Kappa inter-rater reliability estimates showed the HoNOSI to have *Almost Perfect* inter-rater reliability for the HoNOSI total score. Of the 15 scales, 1 had *Moderate*, 7 had *Substantial*, and 7 had *Almost Perfect* inter-rater reliability, exceeding the COSMIN standard for inter-rater Reliability ($K_w \geq 0.7$).

Conclusion

In summary, the NOCC Strategic Review highlighted the need for an instrument such as HoNOSI. The concurrent validity study findings were positive and acceptable. This study provides evidence of inter-rater reliability. The study is limited by the inclusion of some participants with outcome measurement but not infant mental health backgrounds. Future research will address this issue. Nonetheless, the current findings combined with the previous on face and concurrent validity studies, support release of the HoNOSI.

1. BACKGROUND

The Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA) was developed in the UK by a team led by Dr Simon Gowers and Prof Richard Harrington in the late 1990s (S. G. Gowers, R. C. Harrington, A. Whitton, et al., 1999). The aim was to have an instrument that was brief, reliable and valid enough that it could be used routinely in child and adolescent mental health services (CAMHS) to ascertain both a snapshot of children and adolescent's level of symptomatology as well as change over time. The first HoNOSCA paper was published in 1999 and since then HoNOSCA has been used widely (Adamis, Giannakopoulou, Konstantopoulou, et al., 2011; Bilenberg, 2003; Brann and Coleman, 2010; Duffy and Skeldon, 2014; S.G. Gowers, R.C. Harrington, A. Whitton, et al., 1999; Hanssen-Bauer, Gowers, Aalen, et al., 2007; Kisely, Campbell, Crossman, et al., 2007; Lesinskiene, Senina and Ranceva, 2007).

Australian CAMHS trialled HoNOSCA in 1999 and by 2003, it was part of the National Outcomes and Casemix Collection (NOCC). The NOCC is the national specification of a set of standardised measures of consumer symptoms and functioning and a protocol for their collection at key points of care within public specialised clinical mental health services (Technical Specifications Drafting Group. Information Strategy Committee. Australian Health Ministers Advisory Council. National Mental Health Working Group, 2003). All Australian State and Territory Governments States and Territories began implementing the NOCC in 2000-01; by 2009, all jurisdictions were collecting and reporting the NOCC (Burgess, Pirkis and Coombs, 2015).

The Child and Adolescent Mental Health Information Development Expert Advisory Panel (CAMHIDEAP) recognised the need for related instruments for young children (the under 4 year old age group). This was identified as a gap in the measures specified for collection in the NOCC protocols. Using a set of parameters for identifying suitable instruments, CAMHIDEAP undertook an extensive literature review, canvassing both clinician and parent instruments. No suitable measures, covering the entire age range and the key domains for measurement, were able to be identified in the review of the literature (Annex 2). It was agreed that another option was therefore to create an instrument similar to HoNOSCA, but suitable for the 0-4 year old age band.

The Health of the Nation Outcome Scale for Infants (HoNOSI) (Appendix 1) has been developed by the Australian Child and Adolescent Mental Health Information Development Expert Advisory Panel (CAMHIDEAP) as a routine outcome measure for clinicians working with the emotional and social well-being of children in the 0-47 month age group. With a collaborative relationship already established during an international examination of HoNOSCA (Hanssen-Bauer, Gowers, Aalen, et al., 2007), the lead initiative in the HoNOSCI (now known as the HoNOSI) development process occurred from Dr Sally Merry of New Zealand, with in principle support from key figures of the collaboration from the UK, Denmark, Norway, as well as Australia. With the initial ideas scoped by Dr Merry, the CAMHIDEAP facilitated further development of the HoNOSI. The working group has included Dr Peter Brann, Dr Elisabeth Hoehn, Ms Margaret Hoyland, Dr Anne Sved Williams, Dr Nick Kowalenko, Dr Tim Coombs and Ms Rosemary Dickson.

The National Outcomes and Casemix Collection (NOCC) Strategic Directions 2014 – 2024 report (National Mental Health Information Development Expert Advisory Panel, 2013) on the implementation and future directions for the NOCC identified a measurement gap for infants and pre-schoolers. The CAMHIDEAP and a project working group subsequently developed the HoNOSI, based upon the structure of the HoNOSCA, in order to fulfil this need. The decision to utilise the HoNOSCA as the foundation, rather than develop a completely new instrument, was partly informed by the desire to address this national need, the absence of suitable alternative options and the pragmatics of having an instrument that was constructed separately by all of the Australian state and territory governments for the collection of the HoNOSCA measures.

Underpinning the development of HoNOSI was the need for it to be pragmatic and adequate for use in routine clinical practice. While the instrument may eventually be used outside the mental health system, its target group is the public mental health system (with, hopefully, application to the private mental health sector). The instrument parallels the 15 item structure of the HoNOSCA as all governments in both Australia and New Zealand have existing mental health information system infrastructure that can accommodate the collection of data via that form.

Following initial work to develop a draft HoNOSI, CAMHIDEAP undertook consultations with a range of CAMHS and infant mental health clinicians. This consultation identified several areas which required further development. Over recent months, the working group has continued to modify the measure – aiming for clarity and consistency. After an initial survey on the usefulness and relevance of the HoNOSI, the CAMHIDEAP Working Group identified the need for face validity testing to inform further development. HoNOSI will eventually be used by staff in CAMHS and specialist infant mental health services and staff in other related organisations, such as parenting organisations. It was therefore important to ensure that its language and content are meaningful to those clinicians and that the tool appears to be measuring what it is intended to measure. This face validity phase, with feedback on the draft measure provided by focus groups of clinicians, resulted in some further modifications to the draft measure (Annex 1).

Face validity testing (Annex 1) showed that the HoNOSI fulfilled a much needed gap in infant mental health outcome measurement for the 0-47 month age group as no suitable instrument previously existed. Overall, regarding the overarching question on the purpose and usefulness of the HoNOSI and whether it fulfils its objectives, concluding thoughts were unanimously positive in responses from all 6 out of the 6 face-to-face focus groups. These focus groups were conducted in all participating jurisdictions across Australia during face validity testing. Participants expressed a want and need to use it and welcomed the opportunity to have input and to make suggested improvements from a user perspective to inform the next stages of development. Annex 1 refers to the report on the face validity testing phase of development. Following face validity testing, the CAMHIDEAP Working Group identified the need for field testing to test selected psychometric properties of the HoNOSI.

1.1 Aim

To test the concurrent validity and inter-rater reliability of the HoNOSI.

1.2 Psychometric Properties to be tested

The COnsensus based Standards for the selection of health status Measurement INstruments (COSMIN; https://www.cosmin.nl/) (Amsterdam Public Health, 2018) initiative was developed to provide guidance on the selection of high quality patient-reported outcome measures to clinical and research applications (Prinsen, Mokkink, Bouter, et al., 2018). It includes: (i) a taxonomy and definitions of measurement properties (Mokkink, Terwee, Patrick, et al., 2010), checklists for assessing the methodological quality of an individual study of measurement properties (Mokkink, de Vet, Prinsen, et al., 2018) and criteria for good measurement properties, against which to evaluate the results of an individual study (Prinsen, Mokkink, Bouter, et al., 2018).

This framework has been applied to the suite of Health of the Nation Outcome Scales clinician-reported measures (M. Harris, Sparti, Scheurer, et al., 2017; M. G. Harris, Sparti, Scheurer, et al., 2018). It is used here to inform evidence regarding the psychometric properties of the HoNOSI.

The COSMIN framework is comprehensive with respect to psychometric properties that should be evaluated. The HoNOSI field trial was designed to examine two of those psychometric properties:

1. **Concurrent validity** – how well HoNOSI ratings correlate with other measures of similar constructs; and,
2. **Inter-rater reliability** – exploring the degree of agreement among different raters' ratings of the same case vignette.

1.3 Methods Common to Study One and Study Two

1.3.1 Recruitment

The CAMHIDEAP nominated key clinicians across a range of jurisdictions who were engaged in providing mental health services to infants and pre-schoolers. The feasibility of using these services was established through exploring the numbers of infants seen and the types of measures routinely collected for infants and adults. Most of these clinicians had previously been involved in the face validity study. Services from jurisdictions previously participating in the face validity study were invited to participate; Queensland, NSW, Victoria, South Australia and Western Australia.

1.3.2 Materials

After establishing a service coordinator for each of the nine sites, an *Information for Co-ordinators* summary was provided. Co-ordinators explained the study and asked participants to sign a consent form in order to be able to participate. Along with an explanation from the site co-ordinator, the co-ordinator highlighted the rationale, background, aims and underscored that the clinician's information would remain confidential and be analysed in aggregate, anonymous form only.

Participants were also provided with an information sheet which they were asked to read before signing the consent form to participate. Copies of the study protocol were included with the study material for co-ordinators' and participants' reference. Participants were encouraged to ask questions at any time of the site co-ordinator or the project co-ordinator.

Information Sheets and Consent Forms had site specific elements as each Ethics Committee and/or Research and Governance Office had unique requirements. The Information Sheet and Consent Form were stapled together and printed with blank pages inserted where necessary. Signed consent forms and completed ratings were returned to the "HoNOSI box". All sites were provided with a return address label to be adhered to the HoNOSI box, for collection by courier. For smaller bundles, a self-addressed express post envelope was provided for posting the signed consent forms and completed ratings back to the Health Education and Training Institute (HETI).

The HoNOSI (Appendix 1) has 15 Scales that are each rated on a 5-point rating scale:

0 - No problem.
1 - Minor problem requiring no formal action.
2 - Mild problem.
3 - Problem of moderate severity.
4 - Severe to very severe problem.
9 - Not known or not applicable.

A separate rating value of 9 can be recorded for each Scale where information is *Not known/not applicable*. The total score is calculated by summing the first 13 scales. Missing data is treated as zero in calculating totals. The HoNOSI is comprised of 15 scales, as shown in Table 1.

Table 1: *HoNOSI Scales*

Scale	Scale Name
Scale 1	Problems with disruptive behaviour/irritability/under controlled emotional regulation
Scale 2	Problems with activity levels, joint and/or sustained attention
Scale 3	Non accidental self-injury or lack of self-protective behaviours
Scale 4	Problems with feeding and eating behaviour
Scale 5	Problems with developmental delays
Scale 6	Problems with physical illness or disability
Scale 7	Problems associated with regulation and integration of sensory processing
Scale 8	Problems associated with sleep
Scale 9	Problems with emotional and related symptoms or over-controlled emotional regulation
Scale 10	Problems with social reciprocity
Scale 11	Problems with age appropriate self-care and environmental exploration
Scale 12	Problems with family life and relationships
Scale 13	Problems with attending care, education and socialisation settings
Scale 14	Problems with knowledge or understanding about the nature of the infant's difficulties
Scale 15	Problems with lack of information, understanding about services, or managing the infant's difficulties

1.3.3 Ethics

Ethics and site-specific approval for both studies were obtained from each participating jurisdiction; NSW, Queensland, Victoria, South Australia and Western Australia. As this study was asking clinicians to complete outcomes scales, Low-Negligible Risk Ethics approvals were sought from the Human Research Ethics Committees that govern the areas in which the relevant services were located, as well as site specific approval from the respective Research and Governance Office in each participating site within each jurisdiction.

1.3.4 Consent

The concurrent validity study required clinicians to sign consent forms and also to complete outcome measures for infants referred to their service. Returned ratings did not identify the clinician nor the infant. In the rare event that the clinician returned a completed rating scale(s) without having completed a consent form, consent was implied via the participant's return of the outcome rating. No information was sought from children, infants or parents. The inter-rater reliability study required clinicians to complete the HoNOSI for case vignettes.

The following section describes the field trial in two parts. *Study One* reports field testing of concurrent validity of the HoNOSI, with clinical cases as subjects. *Study Two* reports field testing of inter-rater reliability of the HoNOSI using case vignettes.

2. STUDY ONE: CONCURRENT VALIDITY

In order to test concurrent validity, the HoNOSI needs to be compared to other measure(s) or a *Gold Standard* which measures similar constructs. As there is no such measure currently being routinely collected in jurisdictions (refer to Annex 2 for a review of the literature and a review of infant outcomes measures), the working group determined that the best comparison feasibly available is to test the HoNOSI (Appendix 1) against the Parent-Infant Relationship Global Assessment Scale (PIR-GAS; Appendix 2), as it is routinely collected in one jurisdiction (Queensland). To partially solve the issue of these two instruments measuring different constructs, two simple scales, developed by the project working group - the *Clinical Worry* and *Severity Judgement* ratings were also rated concurrently (Appendix 3), as shown in Figure 1. The closer these sets of scores are to showing similar outcomes when compared against each other, the higher the concurrent validity of the HoNOSI to the Gold Standard (PIR-GAS + Clinical Worry Likert Rating + Severity Judgement Likert Ratings). Concurrent validity was assessed with actual clients in routine clinical care.

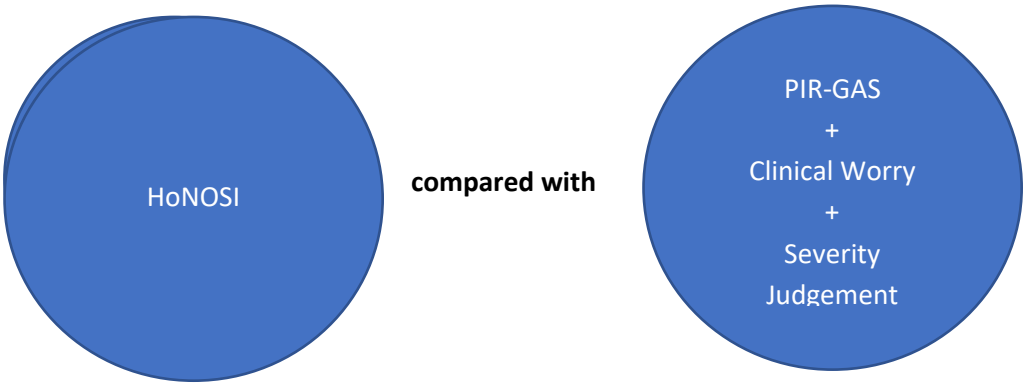

Figure 1: Conceptual Model of Testing Concurrent Validity: HoNOSI Versus PIR-GAS + Clinical Worry + Severity Judgement

2.1 Method

Clinicians completed the HoNOSI, PIR-GAS, Clinical Worry Rating and Severity Judgement Rating for each clinical case. Concurrent validity was to be assessed by comparing the HoNOSI against these ratings; PIR-GAS, Clinical Worry and the Severity Judgement Ratings. Each of these measures is described in more detail below.

The study material was designed as a ready-to-use stapled "booklet", which included, in order, a full version of the HoNOSI, PIR-GAS, Clinical Worry and Severity Judgement Ratings. The co-ordinators were asked to keep these and the HoNOSI box in a central place for easy access to and return of the ratings booklets.

2.1.1 Sample Size

As effect size is unknown, determining the sample size required for statistical analysis was guided by examining the literature in other similar studies. In terms of measuring concurrent validity, similar studies ranged in size from 51 (Harnett, Loxton, Sadler, et al., 2005) to 248 (Yates, Garralda and Higginson, 1999) participants. These studies compared the HoNOSCA to other indicators. From the literature, one of the most similar study types examined the correlation between HoNOS and the Global Assessment of Functioning (GAF), which assessed 100 consecutive outpatients (Browne, Doran and McGauran, 2014). Another similar study included 79 patients (Parker, O'Donnell, Hadzi-Pavlovic, et al., 2002). A reliability and validity study assessed concurrent validity (n=151) of the HoNOS and found correlations were high for most ratings (Amin, Singh, Croudace, et al., 1999). That study compared the HoNOS to other scales, including the GAF. A previous AMHOCN report on the development of a mental health carer experience service provision measure used 41 cases in an analysis of test retest reliability (Australian Mental Health Outcomes and Classification Network, 2018). Therefore, in terms of literature to date and feasibility, the aim of this study was to recruit clinicians to rate 60-80 cases.

2.1.2 Measures

The HoNOSI, as outlined above, was used in the concurrent validity study. The PIR-GAS (Appendix 2) is a measure of the quality of the infant-parent relationship (The Washington Institute and the Washington State Mental Health Division, 2012). Clinicians assess the intensity, frequency and duration of difficulties on a 100 point rating scale that ranges from 1 *Documented Maltreatment* to 100 *Well Adapted*. PIR-GAS ratings are usually reported in deciles (e.g., 1-10, 11-20, etc.) with a higher order classification of 3 levels: 1 – 40 *Disordered Relationship*, 41-80 *Features of Disordered Relationship* and 81-100 *Adapted Relationship*.

The Clinician Worry Rating (Appendix 3) is 7 point rating scale, developed by the HoNOSI project working group, asks the clinician to rate: "Overall, how concerned are you about this infant?"

The Severity Judgement Rating (Appendix 3) is also a 7 point rating scale, developed by the HoNOSI project working group, which asks the clinician to rate: "In your clinical judgement, how severe do you consider the infant's overall social and emotional problems?"

Both the Clinician Worry Rating and the Severity Judgement Rating scales were designed to be unidirectional; to have 0 at the *Not worried/No problem* end of the Likert scale and *Extreme* and *Severe* rating at the high end of the scale.

2.1.3 Data Analysis

Statistical analyses were performed using SPSS Version 24 (IBM Corp, 2016) and Stata Version 14.2. (StataCorp., 2015).

2.1.4 Data Collection of Clinical Cases

All but 5 cases met the age range criteria of 0-47 months for inclusion in the analysis. These cases were excluded as the age of the infant was either out of range (0-47 months), not provided, or was unable to be verified.

2.2 Descriptive Statistics - Clinical Cases

The Study 1 analysis dataset consisted of 108 completed clinical cases. These were returned by participating clinicians from 4 jurisdictions across Australia, including 5 participating services. Data collection of clinical cases occurred over the period June 2017 through May 2018.

As per ethics approval requirements, the responses and the identity of participants was kept confidential. Participants were informed of this and that their responses would be analysed in aggregate, anonymous form only. Therefore, it was not possible to determine the exact number of participating clinicians, as each clinician could have rated one or more infants. The estimated number of clinicians participating at each site was based on the clinician's reported professional background and their reported years of experience. Table 2 shows a breakdown of the estimated

number of clinicians and the number of clinical cases per site. It is estimated that a minimum of 26 clinicians participated in rating clinical cases.

Table 2: *Concurrent validity study sites by number of clinicians and infants rated*

Jurisdiction	Site	Clinicians (min estimate) *n*	Infants *n*
Qld	Children's Health Queensland Hospital and Health Service	11	55
NSW	Perinatal and Infant Mental Health/Child and Youth Mental Health Service, Northern Sydney LHD	8	29
NSW	Perinatal and Infant Mental Health Service, Western Sydney LHD	1	6
Vic	Alfred Health (CYMHS)	3	7
SA	Women's and Children's Health Network (CAMHS)	3	11
Total		26	108

2.2.1 Profession Characteristics

Table 3 shows the estimated number of clinicians and the number of ratings completed by profession type. Over half of the clinicians were either Psychologists or Social Workers and these two profession types completed approximately two-thirds of all clinical ratings.

Table 3: *Infants rated by clinician's profession type*

Profession	Clinicians		Infants	
	n	%	*n*	%
Psychologist	8	30.8	41	38.0
Social Worker	6	23.1	30	27.8
Psychiatrist	4	15.4	10	9.3
Nurse	3	11.5	15	13.9
Occupational Therapist	1	3.8	7	6.5
Speech Pathologist	1	3.8	1	0.9
Unknown	3	11.5	4	3.7
Total	26	100.0	108	100.0

2.2.3 Clinical Experience

Table 4 shows the estimated number of clinicians and the number of ratings completed by the clinicians' years of experience. Over 61% of clinicians rating the clinical cases had over 5 years' experience and of those, over 23% had clinical experience of over 10 years; these clinicians rated over 75% of the clinical cases.

Table 4: *Infants rated by clinicians' years of experience*

Clinical Experience	Clinicians		Infants	
	n	%	*n*	%
<2 Years	5	19.2	15	13.9
2-5 years	4	15.4	9	8.3
5-10 years	10	38.5	47	43.5
10+ years	6	23.1	35	32.4
Unknown	1	3.8	2	1.9
Total	26	100.0	108	100.0

2.2.4 Infant Characteristics

Basic demographic data were collected regarding the age and sex of the infant. These infant characteristics, along with their ratings are described below. Table 5 shows the age distribution. It can be seen that there were slightly more male (52.8%) than female infants (47.2%). The age distributions differed; male infants were somewhat older than female infants, with median ages of 16 and 10 months, respectively.

Table 5: *Age Distribution of Clinical Cases (n = 108)*

Sex	*n*	mean	*SD*	min	max	percentile				
						10th	25th	50th	75th	90th
Male	57	18.3	14.0	1	47	3	5	16	30	38
Female	51	16.3	13.9	1	47	2	5	10	25	41
Total	108	17.3	14.0	1	47	2	5	15	26	41

2.2.5 Distribution of PIR-GAS Ratings

Table 6 shows the distribution of PIR-GAS ratings; there were 4 infants who had no PIR-GAS rating recorded. It can be seen that 8.7% of infants were classified in terms of *Adapted Relationship* (PIR-GAS rating 81-100), 49.0% *Features of Disordered Relationship* (PIR-GAS rating 41-80) and 42.3% *Disordered Relationship* (PIR-GAS rating 1-40).

Table 6: *Distribution of PIR-GAS ratings (n = 104)*

PIR-GAS Ratings for Clinical Cases	%
1-10 Documented Maltreatment	0.0
11-20 Grossly Impaired	1.0
21-30 Severely Disordered	11.5
31-40 Disordered	29.8
41-50 Disturbed	15.4
51-60 Distressed	16.3
61-70 Significantly Perturbed	8.7
71-80 Perturbed	8.7
81-90 Adapted	5.8
91-100 Well Adapted	2.9
Total[1]	100.0

[1]There were 4 non-responses to the PIR-GAS

2.2.6 Distribution of Clinical Worry and Severity Judgement Ratings

Table 7 shows the distributions of clinicians' Clinical Worry and Severity Judgement ratings. The majority of infants were rated as either a "3" or a "4" in terms of Clinical Worry (51.0%) and Severity Judgement (43.6%) on the respective 7-point rating scales.

Table 7: *Distribution of Clinical Worry and Severity Judgement ratings (n = 108)*

Clinical Worry (%)		Severity Judgement (%)	
Not Concerned-0	10.2	No Problem-0	15.7
1	9.3	1	11.1
2	12.0	2	16.7
3	26.9	3	23.2
4	24.1	4	20.4
5	12.0	5	8.3
Extremely Concerned-6	5.6	Extremely Severe Problem-6	4.6
Total	100.0	Total	100.0

2.2.7 Distribution of HoNOSI Ratings

The frequency distribution of severity ratings for each of the 15 HoNOSI Scales is presented in Figure 2. Within each Scale, each bar represents the overall percentage of cases rated per the 5-point severity rating. The frequency distribution of HoNOSI Ratings for Clinical Cases can be found in Appendix 4.

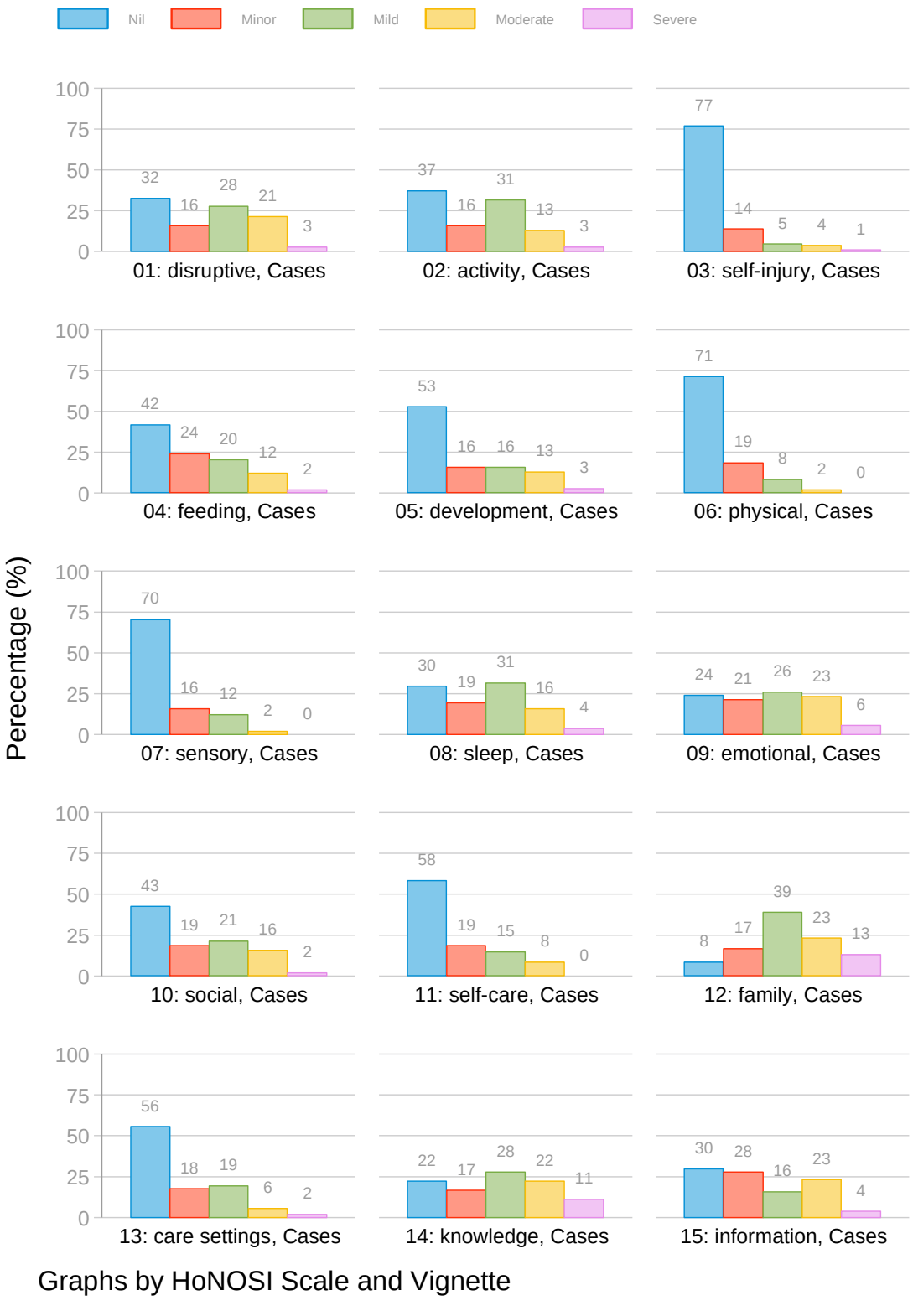

Figure 2: Ratings for the 15 HoNOSI Scales

No infants received a HoNOSI rating of 9 (Not known/Not Applicable), nor were there any missing ratings on clinical cases (Appendix 4). Of the 108 clinical cases, all 5 rating points were used for 12 of the 15 HoNOSI Scales; the most severe rating of 4 was not used for Scales 6 *Problems with physical illness or disability*, 7 *Problems associated with regulation and integration of sensory processing* and 11 *Problems with age appropriate self-care and environmental exploration*. For five HoNOSI Scales; *3 Non accidental self-injury or lack of self-protective behaviours*, 5 *Problems with developmental delays*, 6 *Problems with physical illness or disability*, 7 *Problems associated with regulation and integration of sensory processing* and 11 *Problems with age appropriate self-care and environmental exploration*, over half all of ratings were rated 0, indicating *no problems/issues*.

Ratings on Scale 1 *Problems with disruptive behaviour/irritability/under controlled emotional regulation* were spread across all the rating points, although there were only a few severe ratings provided. In contrast, Scale 3 *Non accidental self-injury or lack of self-protective behaviours* had only a small percentage with clinically significant ratings. The overwhelming majority received *no problem* or *minor problem* ratings.

Clinicians typically classify ratings on the HoNOS family of measures as "clinically significant" (Burgess, Trauer, Coombs, et al., 2009), if a problem area is rated as either a *Mild*, *Moderate* or *Severe to Very Severe* problem (i.e., a rating of 2, 3 or 4). Using that classification with respect to the 108 clinical cases, 75% were rated as having "clinically significant" problems with Scale 12 *Problems with family life and relationships*, 55% with Scale 9 *Problems with emotional and related symptoms or over-controlled emotional regulation*, 52% with Scale 1 *Problems with disruptive behaviour/irritability/under controlled emotional regulation* and 51% with Scale 8 *Problems associated with sleep*. The least frequently rated areas with "clinically significant" problems were Scale 3 with 9% *Non accidental self-injury or lack of self-protective behaviours*, Scale 6 with 10% *Problems with physical illness or disability* and Scale 7, 14% *Problems associated with regulation and integration of sensory processing*.

"Clinically significant" problems were also found for 61% on Scale 14 *Problems with knowledge or understanding about the nature of the infant's difficulties* (61%) and for 43% on Scale 15 *Problems with lack of information, understanding about services, or managing the infant's difficulties* (43%).

2.2.8 Distribution of HoNOSI Total Scores for Clinical Cases

The first 13 HoNOSI scales concern the infant's mental health status and a total HoNOSI score can be derived by summing the ratings of these Scales. The remaining two HoNOSI Scales, 14 *Problems with knowledge or understanding about the nature of the infant's difficulties* and 15 *Problems with lack of information, understanding about services, or managing the infant's difficulties*, do not assess the infant's mental health status and are not used in calculations of the HoNOSI total score.

Table 8 and Figure 3 present summary descriptive statistics and the distribution of HoNOSI total scores. HoNOSI total scores ranged from 0 through 42 (Table 8), with a mean and median of 14.0 and an interquartile-range of 12 points (i.e., the middle 50% of total scores were within the range 7 through 19).

The shape of the distribution of HoNOSI total scores shows small but statistically significant positive skewness (0.47). A statistical test of the overall distribution, however, was not statistically significant (Shapiro-Francia *W' = 0.98, p = 0.16)* indicating a normal distribution of severity scores.

Collins et al (2016) define floor and ceiling effects for measures reporting total scores as "the percentage of respondents with the lowest possible score (floor effects) and the highest possible score (ceiling effects)" (Collins, Prinsen, Christensen, et al., 2016, p. 1319). Floor and ceiling effects are not considered statistically significant if less than 15% of subjects score the lowest or the highest possible score. There was no evidence of these effects in the HoNOSI total scores.

Table 8: *Distribution of HoNOSI Total Score in Clinical Cases (n = 108)*

	n	*m*	*sd*	min	max	Percentile				
						10th	25th	50th	75th	90th
Cases	108	14.0	8.8	0	42	3	7	14	19	27

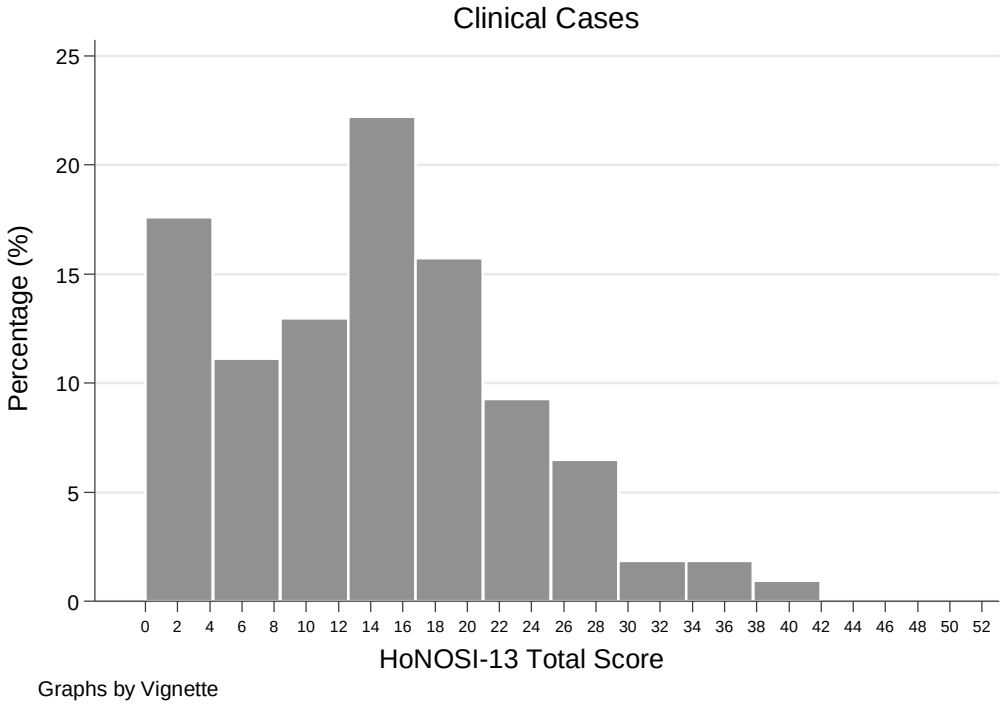

Figure 3: Distribution of HoNOSI Total Scores for Clinical Cases

In COSMIN, internal consistency is defined as "The degree of interrelatedness among the items" (Mokkink, Terwee, Patrick, et al., 2010, p. 743) and the minimum "acceptable" COSMIN standard for internal consistency is 0.70 (Prinsen, Mokkink, Bouter, et al., 2018). The level of internal consistency of the 13 Scales comprising the total score, as measured by Cronbach's α, is 0.87. This exceeds the COSMIN threshold.

A HoNOSI “item severity structure” index was derived using the method described by Gowers et al (1999) (S. G. Gowers, R. C. Harrington, A. Whitton, et al., 1999) with respect to the HoNOSCA. Table 9 shows the item severity structure index by the mean HoNOSI total score. As was found by Gowers et al. with the HoNOSCA, the mean HoNOSI total score increases with the number of individual HoNOSI Scale ratings that are “clinically significant”. It is important to note that more than 80% of clinical cases had at least one HoNOSI Scale problem area rated as “clinically significant”. This finding suggests that the overall clinical severity of these 108 cases is likely representative of very young consumers seen in specialised public sector mental health services.

Table 9: *HoNOSI item severity structure by HoNOSI total score (n = 108)*

HoNOSI Severity index	n	%	Cumulative %	HoNOSI (mean)
Two or more ratings of 4	9	8.3	8.3	27.0
Only one rating of 4	12	11.1	19.4	17.7
No 4 ratings, any ratings of 3	42	38.9	58.3	17.4
No 3 or 4 ratings, any ratings of 2	28	25.9	84.3	10.0
No 2-4 ratings, any ratings of 1	11	10.2	94.4	4.1
0 ratings only	6	5.6	100	0
Total	108	100	100	14.0

2.3 Concurrent Validity Analysis of Clinical Case Data

In COSMIN, criterion validity is defined as "The degree to which the scores on [the] instrument are an adequate reflection of a "gold standard" (Mokkink, Terwee, Patrick, et al., 2010, p. 743). Concurrent validity is considered a part of criterion validity. Concurrent validity is considered "adequate" where the measure correlates with a "gold standard" of at least 0.70 (Mokkink, Prinsen, Patrick, et al., 2018).

If the HoNOSI has "adequate" concurrent validity, then:

- statistically significant *negative* correlation coefficients of at least -0.70 are hypothesized with the PIR-GAS clinical ratings as higher clinical ratings represent a positive status, and;
- statistically significant *positive* correlation coefficients of at least 0.70 are hypothesized with the Clinical Worry and Severity Judgment as higher clinical ratings represents a negative status.

Spearman's rank order correlation was used to test the concurrent validity of the HoNOSI total score with the PIR-GAS, Clinical Worry and Severity Judgement Scales and are presented in Table 10. A comprehensive inter-correlation analysis of the 15 individual HoNOSI Scales (i.e., Spearman rank order correlation coefficients among all Scales) is at Appendix 5. They show that all 15 HoNOSI Scale correlations with the PIR-GAS, Clinical Worry and Severity Judgement ratings are statistically significant (*p<0.001)* with the one exception of Scale 6 correlated against the Clinical Worry rating, which is also statistically significant at a lower threshold (*p<0.05*). The HoNOSI total score correlations summary table (table 10 below) shows that the three validity measures exceed the COSMIN threshold.

It is also important to note that the three concurrent validity measures are highly statistically inter-correlated (*p<0.001*): PIR-GAS with Clinical Worry, r_s = -0.81; PIR-GAS with Severity Judgment, r_s = -0.76; and Clinical Worry with Severity Judgement, r_s =0.82. Details are shown in Table 10. Correlation analyses of the 15 individual HoNOSI Scales with these three ratings can be found in Appendix 6.

Table 10: *Correlation of HoNOSI Total Score with PIR-GAS, Clinical Worry and Severity Judgement Ratings*

Spearman's rho Correlation Coefficient	PIR-GAS (n=104)	Clinical Worry (n=108)	Severity Judgement (n=108)	HoNOSI Total Score (n=108)
PIR-GAS	1.00			
Clinical Worry	-0.81	1.00		
Severity Judgement	-0.76	0.81	1.00	
HoNOSI Total Score	-0.73	0.77	0.85	1.00

[1]All correlations were statistically significant (*p<0.001)*

2.4 Discussion

Study 1 was designed specifically to establish the level of evidence of concurrent validity with respect to the 15 HoNOSI Scales and the HoNOSI total severity score. It also evaluated a number of important psychometric measurement properties.

The clinical case sample characteristics suggest that these 108 very young consumers are likely to be representative of the target population within specialised public sector mental health services. Using the COSMIN criteria, there is evidence for HoNOSI having "adequate" concurrent validity, as assessed by correlations with the PIR-GAS, Clinical Worry and Severity Judgement Scales. There is also evidence for other psychometric properties including the internal consistency and distribution of the HoNOSI total score. The level of internal consistency of the 13 Scales comprising the total score, as measured by Cronbach's shows good internal consistency, exceeding the COSMIN threshold. Notwithstanding those positive findings, the field trial also identified issues requiring further consideration.

The most severe rating of 4 was not used for three of the 15 Scales: Scale 6 *Problems with physical illness or disability*, Scale 7 *Problems associated with regulation and integration of sensory processing,* nor Scale 11 *Problems with age appropriate self-care and environmental exploration*. It could be in this sample of 108 infants, there that were no cases with *Severe to very serve problem* for the HoNOSI problem areas. Alternatively, it could be the description and/or interpretation in the glossary for these Scales means that it is unlikely that a rating of 4 would be used. Future work could further explore these particular Scales in another sample and/or closer examination of these three glossary and ratings descriptions to further refine the scales and ensure full utilisation of the range of possible scores from *0 - No problem* to *4- Severe to very severe problem.*

It should also be noted that the evaluation of concurrent validity was based on three, highly related measures. Future work could explore HoNOSI Scale validity with respect to other domains and consumer attributes including the specific nature of presenting problems and diagnosis categories.

3. STUDY TWO: INTER-RATER RELIABILITY

In order to test inter-rater reliability of the HoNOSI, multiple raters (clinicians) independently completed the HoNOSI, on the same 5 case vignettes. A copy of the 5 vignettes used in this study can be found in Appendix 13.

3.1 Method

Initially, a pilot study with three vignettes was conducted by a panel of CAMHIDEAP working group members, along with individuals with expertise in infant mental health or mental health measurement (n=11), of which full details can be found in Appendix 7. This pilot study identified poor reliability with substantial overlap between the intended levels of severity and upon further investigation, a likely explanatory factor was a lack of variability between the three vignettes used for the pilot. To that end, the three original vignettes were rewritten by one member of the working group to ensure a spread of severity over each of the domains. Along with refining the three original vignettes, two additional newly created vignettes were added to the test collection, expanding the pool of vignettes for testing from three to a total of five vignettes to be completed by each rater in the current study. As described in the results section, these two approaches, applied together, were shown to successfully resolve any lack of variability issues.

For the current study, participants with expertise in infant mental health or mental health measurement, were asked to rate these vignettes. Ethics and Research and Governance approval was obtained from the Children's Health Queensland Hospital and Health Service (CHQHHS) for participants from the Queensland Centre for Perinatal and Infant Mental Health and the Child and Youth Mental Health Service to participate in the study. Likewise, ethics approval was obtained from Eastern Health, Victoria in order to be able to include participants from the Child and Youth Mental Health Service. A total of 45 people participated in the study; 26 from Queensland and 19 from Victoria. All participants received a brief online training session on the measure (Brann and Coombs, 2020), the rating rules and inclusion and exclusion criteria. Upon providing consent via the online consent form, the five vignettes were independently rated online. Those making the ratings were presented with the vignette and then the HoNOSI glossary from which the rater selected their preferred rating. Results were analysed using STATA statistical software Version 16.1 (StataCorp LLC, 2020). Inter-rater reliability was assessed using quadratic weighted kappa. A copy of the vignettes is found in Appendix 13.

The HoNOSI (Commonwealth of Australia for and on behalf of the Australian Mental Health Outcomes and Classification Network, 2020) is a 15-item set of scales rated as:

0 - No problem.
1 - Minor problem requiring no formal action.
2 - Mild problem.
3 - Problem of moderate severity.
4 - Severe to very severe problem.
9 - Not known or not applicable.

The total score is calculated by summing the first 13 scales which relate to infant mental health status. Missing data is treated as zero in calculating totals. As with the HoNOS family of measures,

clinicians typically classify ratings as 'clinically significant' (Burgess, Trauer, Coombs, et al., 2009), if a problem area is rated as either a *Mild*, *Moderate* or *Severe to Very Severe* problem (i.e., a rating of 2, 3 or 4). A list of HoNOSI scales can be found at Appendix 1.

The strength of agreement between raters were measured against the standards outlined by Landis and Koch (1977) (Landis and Koch, 1977, p. 165) and by the Consensus-based Standards for the selection of health status Measurement Instruments (COSMIN) (Amsterdam Public Health, 2018). In contrast, the COSMIN initiative (Amsterdam Public Health, 2018) was developed to provide guidance on the selection of outcome measures to clinical and research applications (Prinsen, Mokkink, Bouter, et al., 2018). This includes providing a methodology for assessing the interrater reliability of measures (Amsterdam Public Health, 2018; L.B. Mokkink, Prinsen, Patrick, et al., 2018). It comprises a taxonomy and definitions of measurement properties (L. B. Mokkink, Terwee, Patrick, et al., 2010), checklists for assessing the methodological quality of measurement properties (L. B. Mokkink, de Vet, Prinsen, et al., 2018) and criteria for good measurement properties, against which to evaluate study results (Prinsen, Mokkink, Bouter, et al., 2018). In COSMIN, for measures that are constructed for ratings at an ordinal level of measurement (i.e., the rating categories indicate the relative ordering of "clinical severity"), the "gold standard" for reporting is the weighted kappa (L.B. Mokkink, Prinsen, Patrick, et al., 2018) where reliability is "sufficient" if the weighted Kappa, $K_w \geq 0.7$).

3.2 Results

Of the 45 raters, 39 completed all five vignettes with no missing values. Five raters had a missing rating for one scale with one vignette and one rater had a missing rating for one scale with two vignettes. No HoNOSI Scale ratings were rated as 9 *Not known/not applicable*. Table 11 and Figure 4 show descriptive statistics and the distribution of the HoNOSI total scores for each of the five case vignettes.

Table 11: *Distribution of HoNOSI Total Score for the 5 Vignettes*

Vignette #	n	mean	sd	min	max	Percentile				
						10th	25th	50th	75th	90th
1 – Levi (v1)	45	20.7	3.9	14	31	17	18	20	22	27
2 – Lily (v2)	45	17.8	4.4	12	31	13	15	17	20	23
3 – Chloe (v3)	45	10.3	3.6	2	22	6	8	10	12	13
4 – Sandy (v4)	45	31.1	4.0	26	42	26	28	31	33	35
5 – Sally (v5)	45	37.6	4.3	27	46	31	36	37	40	43

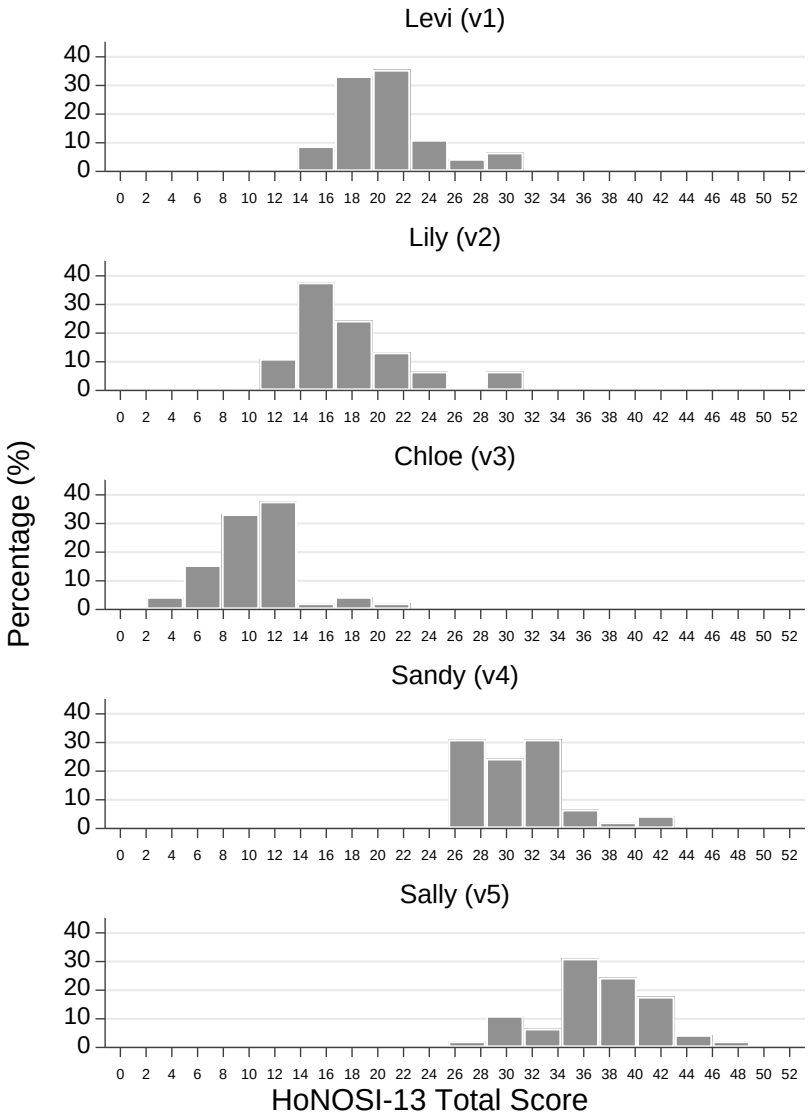

Figure 4: Distribution of HoNOSI Scores for the 5 Vignettes and the HoNOSI total score

As shown in Figure 4, the vignette scores have been used to show a variety of clinical presentations. All of the HoNOSI scores (0,1,2,3,4) have been utilised for each of the 15 scales for at least one of the vignettes. This demonstrates that the ranges of possible scores are useful in a variety of clinical cases.

Inter-rater reliability was assessed using quadratic weighted Kappa (K_w) estimates. Table 12 shows that the Inter-rater reliability of the HoNOSI as rated by the 45 raters who completed ratings on five case vignettes shows *Almost perfect* Inter-rater reliability on the HoNOSI total score (K_w=0.85, 95%CI=0.70-1.00) according to the benchmarks outlined by Landis and Koch (Landis and Koch, 1977). It also exceeds the COSMIN standard for acceptable Inter-rater reliability ($K_w \geq 0.7$) (Prinsen, Mokkink, Bouter, et al., 2018, p. 1157).

Table 12: *Inter-rater reliability quadratic weighted kappa estimates for vignettes*

	HoNOSI Scale	Raters (n)	K_w	K_w 95% CI lower	K_w 95% CI upper	Strength of Agreement
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	45	0.86	0.77	0.94	Almost Perfect
2	Problems with activity levels, joint and/or sustained attention	45	0.41	0.00	0.81	Moderate
3	Non accidental self-injury or lack of self-protective behaviours	44	0.85	0.79	0.92	Almost Perfect
4	Problems with feeding and eating behaviour	45	0.85	0.61	1.00	Almost Perfect
5	Problems with developmental delays	45	0.85	0.76	0.95	Almost Perfect
6	Problems with physical illness or disability	45	0.80	0.57	1.00	Substantial
7	Problems associated with regulation and integration of sensory processing	43	0.69	0.35	1.00	Substantial
8	Problems associated with sleep	45	0.85	0.57	1.00	Almost Perfect
9	Problems with emotional and related symptoms or over-controlled emotional regulation	44	0.80	0.55	1.00	Substantial
10	Problems with social reciprocity	45	0.87	0.76	0.98	Almost Perfect
11	Problems with age appropriate self-care and environmental exploration	45	0.62	0.27	0.98	Substantial
12	Problems with family life and relationships	44	0.81	0.64	0.98	Almost Perfect
13	Problems with attending care, education and socialisation settings	44	0.71	0.27	1.00	Substantial
14	Problems with knowledge or understanding about the nature of the infant's difficulties	45	0.66	0.34	0.97	Substantial
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	44	0.69	0.49	0.88	Substantial
	HoNOSI Total Score	45	0.85	0.70	1.00	Almost Perfect

The Landis and Koch guidelines describe the strength of inter-rater reliability (Landis and Koch, 1977) as follows:

< 0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost Perfect

3.3 Discussion

This study was designed to test the Inter-rater reliability of the HoNOSI. Results show that the level of Inter-rater reliability was *Almost Perfect* when measured against the benchmarks outlined by Landis and Koch (Landis and Koch, 1977, p. 165) and that it also well exceeds the COSMIN standard for Inter-rater reliability (L.B. Mokkink, Prinsen, Patrick, et al., 2018). According to these descriptions, the results of this study (Table 12) show that there is *Moderate, Substantial* and *Almost Perfect* agreement between raters for each of the individual HoNOSI scales and *Almost Perfect* agreement between raters on the HoNOSI total score.

COSMIN guidelines describe what constitutes a *sufficient* level of agreement. Using the COSMIN criteria for good measurement properties, the results of the present study well exceed the criteria for inter-rater reliability (where reliability is rated as either sufficient ($K_w \geq 0.7$), insufficient ($K_w < 0.7$) or indeterminate (where K_w is not reported). Scales 7 *Problems associated with regulation and integration of sensory processing*, 11 *Problems with age appropriate self-care and environmental exploration,* 14 *Problems with knowledge or understanding about the nature of the infant's difficulties* and 15 *Problems with lack of information, understanding about services, or managing the infant's difficulties* were classed as insufficient in terms of reliability (Table 12), although they are close to the arbitrary cut-off ($K_w \geq 0.7$).

It is only Scale 2 *Problems with activity levels, joint and/or sustained attention* with $K_w < 0.5$ that is found in this study to be insufficient according to the COSMIN criteria, although assessed as a *moderate* level of agreement according to the guidelines provided by Landis and Koch (Landis and Koch, 1977). Aside from Scale 2 *Problems with activity levels, joint and/or sustained attention*, which was the scale found to have the least agreement, the other 14 HoNOSI scales show a reasonable amount of variability between vignettes. The study had sufficient power to test the null hypothesis at the IRR estimate of 0.5.

More importantly, some of the raters did not have experience in infant mental health. While they were independent of each other in providing the ratings, the inexperience of some with the target population, may impact on the generalisability of this finding in the field. It could be possible that developmentally informed interpretations of behaviours, such as sustained attention, were compromised. While a very encouraging finding, it will be important for future research to examine the reliability question with a sample where clinical practice dominates.

Vignettes are not the same as clinical practice and the HoNOSI may perform better or worse with real cases, yet vignettes have been found to be a valid tool when measuring the quality of clinical

practice (Peabody, Luck, Glassman, et al., 2004). They remain a key method when wishing to ensure that raters are exposed to exactly the same stimuli.

Future research could explore HoNOSI Scale validity with respect to other domains and consumer attributes including the specific nature of presenting problems and diagnostic categories. There are other psychometric properties (eg., sensitivity to change) yet to be investigated. A face validity study (Australian Mental Health Outcomes and Classification Network, 2016) and concurrent validity field trial (Australian Mental Health Outcomes and Classification Network, 2019) have now been completed. The findings have been sufficiently encouraging to support the implementation of the HoNOSI for use in routine clinical practice with this age group.

REFERENCES

Adamis, D., Giannakopoulou, D. F., Konstantopoulou, A. and Michailides, M. (2011). Translation and standardization of the HoNOSCA (Health of the Nation Outcome Scales for Children and Adolescents) scale in a Greek sample. *Clinical Child Psychology & Psychiatry, 16(4), 567-573.*

Amin, S., Singh, S. P., Croudace, T., Jones, P., Medley, I. and Harrison, G. (1999). Evaluating the Health of the Nation Outcome Scales: Reliability and validity in a three-year follow-up of first-onset psychosis. *British Journal of Psychiatry, 174(5), 399-403.*

Amsterdam Public Health (2018). COSMIN helps you select the most suitable outcome measurement instruments. [Available from: https://www.cosmin.nl/]

Australian Mental Health Outcomes and Classification Network (2016). *Health of the Nation Outcome Scales for Infants: Face Validity Study*.

Australian Mental Health Outcomes and Classification Network (2018). *Development of the Mental Health Carer Experience Survey. An Australian Government Funded Initiative.*

Australian Mental Health Outcomes and Classification Network (2019). *Health of the Nation Outcome Scales for Infants: Concurrent Validity Study.*

Bilenberg, N. M. D. P. (2003). Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA). *European Child & Adolescent Psychiatry, 12(6), 298-302.*

Brann, P., Coleman, G. and Luk, E. (2001). Routine outcome measurement in a child and adolescent mental health service: an evaluation of HoNOSCA. The Health of the Nation Outcome Scales for Children and Adolescents. *Australian & New Zealand Journal of Psychiatry, 35(3), 370-376.*

Brann, P. and Coombs, T. (2020). HoNOSI Training Video. [Available from: https://www.youtube.com/watch?v=G9CKBJ-9Ty8&feature=youtu.be.]

Browne, S., Doran, M. and McGauran, S. (2014). Health of the Nation Outcome Scales (HoNOS): use in an Irish psychiatric outpatient population. *Irish journal of Psychological Medicine, 17(1), 17-19.*

Burgess, P., Pirkis, J. and Coombs, T. (2015). Routine outcome measurement in Australia. *International Review of Psychiatry, 27(4), 264-275.*

Burgess, P., Trauer, T., Coombs, T., McKay, R. and Pirkis, J. (2009). What does 'clinical significance' mean in the context of the Health of the Nation Outcome Scales? *Australasian Psychiatry, 17(2), 141-148.* doi:https://dx.doi.org/10.1080/10398560802460453

Collins, N. J., Prinsen, C. A., Christensen, R., Bartels, E. M., Terwee, C. B. and Roos, E. M. (2016). Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties. *Osteoarthritis and Cartilage, 24, 1317-1329.*

Commonwealth of Australia for and on behalf of the Australian Mental Health Outcomes and Classification Network (2020). *HoNOSI Health of the Nation Outcome Scales for Infants (0-47 months). Glossary. Ver 1.0.*

Duffy, F. and Skeldon, J. (2014). A CAMHS Intensive Treatment Service: clinical outcomes in the first year. *Clinical Child Psychology & Psychiatry, 19(1), 90-99.*

Gowers, S. G., Harrington, R. C., Whitton, A., Beevor, A., Lelliott, P., Jezzard, R. and Wing, J. K. (1999). Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA). *The British Journal of Psyhiatry, 174(5), 428-431.*

Gowers, S. G., Harrington, R. C., Whitton, A., Lelliott, P., Beevor, A., Wing, J. and Jezzard, R. (1999). Brief scale for measuring the outcomes of emotional and behavioural disorders in children. Health of the Nation Outcome Scales for children and Adolescents (HoNOSCA). *British Journal of Psychiatry, 174, 413-416.*

Hanssen-Bauer, K., Gowers, S., Aalen, O. O., Bilenberg, N., Brann, P., Garralda, E., Merry, S. and Heyerdahl, S. (2007). Cross-national reliability of clinician-rated outcome measures in child and adolescent mental health services. *Administration & Policy in Mental Health, 34(6), 513-518.*

Harnett, P. H., Loxton, N. J., Sadler, T., Hides, L. and Baldwin, A. (2005). The Health of the Nation Outcome Scales for Children and Adolescents in an adolescent in-patient sample. *Australian & New Zealand Journal of Psychiatry, 39(3), 129-135.*

Harris, M., Sparti, C., Scheurer, R., Pirkis, J. E., Burgess, P., Coombs, T., Ruud, T., Kisely, S., Hanssen-Bauer, K. and Siqveland, J. (2017). Measurement properties of the Health of the Nation Outcome Scales (HoNOS) family of measures: a systematic review. [Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017057871]

IBM Corp (2016). IBM SPSS Statistics (Version 24).

Kisely, S., Campbell, L. A., Crossman, D., Gleich, S. and Campbell, J. (2007). Are the Health of the Nation Outcome Scales a valid and practical instrument to measure outcomes in North America? A three-site evaluation across Nova Scotia. *Community Mental Health Journal, 43(2), 91-107.*

Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics,* 33(1), 159-174

Lesinskiene, S., Senina, J. and Ranceva, N. (2007). Use of the HoNOSCA scale in the teamwork of inpatient child psychiatry unit. *Journal of Psychiatric & Mental Health Nursing, 14(8), 727-733.*

Mokkink, L. B. (2018). COSMIN Risk of Bias checklist, July 2018. [Available from: https://cosmin.nl/wp-content/uploads/COSMIN_risk-of-bias-checklist_dec-2017.pdf]

Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M. and Terwee, C. B. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of Life Research, 27(5), 1171-1179.*

Mokkink, L. B., Prinsen, C. A. C., Patrick, D. L., Alonso, J., L.M., B., de Vet, H. C. W. and Terwee, C. B. (2018). *COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs). User Manual.*

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M. and de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63(7), 737-745.*

National Mental Health Information Development Expert Advisory Panel (2013). *Mental Health National Outcomes and Casemix Collection: NOCC Strategic Directions 2014 – 2024*.

Parker, G., O'Donnell, M., Hadzi-Pavlovic, D. and Proberts, M. (2002). Assessing Outcome in Community Mental Health Patients: A Comparative Analysis of Measures. *International Journal of Social Psychiatry, 48(1), 11-19.*

Peabody, J. W., Luck, J., Glassman, P., Jain, S., Hansen, J., Spell, M. and Lee, M. (2004). Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Annals of Internal Medicine,* 141(10), 771-780

Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W. and Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research,* 27(5), 1147-1157.doi:10.1007/s11136-018-1798-3

StataCorp LLC (2015). Stata Statistical Software: Release 14.2 USA: College Station, TX: StataCorp LP.

StataCorp LLC (2020). Stata Statistical Software: Version 16.1. USA: College Station, TX: StataCorp LP.

Streiner, D. L., Norman, G. R. and Cairney, J. (2014). *Health Measurement Scales: A Practical Guide to their Development and Use*: Oxford University Press.

The Washington Institute and the Washington State Mental Health Division (2012). Parent-Infant Relationship Global Assessment Scale (PIR-GAS). [Available from: http://depts.washington.edu/washinst/Resources/CGAS/PIRGAS%20Index.htm]

Yates, P., Garralda, M. E. and Higginson, I. (1999). Paddington Complexity Scale and Health of the Nation Outcomes Scales for Children and Adolescents. *British Journal of Psychiatry, 174(5), 417-423.*

BIBLIOGRAPHY

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20, 37-46.*

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70, 213-220*.

D'Agostino, R. B., A. J. Belanger, and R. B. D'Agostino, Jr. (1990). A suggestion for using powerful and informative tests of normality. *American Statistician 44: 316–321.*

Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions*. John Wiley & Sons.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76, 378-382*.

Klein, D. (2018). "KAPPAETC: Stata module to evaluate interrater agreement," *Statistical Software Components S458283*, Boston College Department of Economics.

Royston, P. (1991). sg3.5: Comment on sg3.4 and an improved D'Agostino test. Stata Technical Bulletin 3: 23–24. Reprinted *in Stata Technical Bulletin Reprints, vol. 1, pp. 110–112*. College Station, TX: Stata Press

Appendix 1: Health of the Nation Outcome Scales for Infants (HoNOSI)

HoNOSI
Health of the Nation Outcome Scales for Infants

(0-47 months)

Glossary

Ver 1.0

2020

Comment or questions to: AMHOCN contact@amhocn.org

The Health of the Nation Outcome Scales for Infants (HoNOSI)

Introduction

The early years are foundational for healthy social and emotional wellbeing across the lifespan. Early intervention can have a powerful impact on creating positive outcomes across this developmental trajectory.

The Health of the Nation Outcome Scales for Infants (HoNOSI) has been developed as a routine outcome measure to support clinicians who work to improve the emotional and social well-being of infants and young children in the 0 - 3 year old age group (i.e. 0-47 months). There is a broad range of development within this age band and where appropriate, the HoNOSI attempts to provide guidance for rating both those at the younger and those at the upper end of the age band. At the same time, HoNOSI is built on clinical judgement as to the level of severity to be rated. For further information on developmental stages, refer to:
https://wcwpds.wisc.edu/childdevelopment/resources/CompleteDevelopmentDetails.pdf
https://childdevelopmentinfo.com/ages-stages/#.WCqgZkbD92g
http://raisingchildren.net.au/
http://www.education.vic.gov.au/childhood/parents/health/Pages/default.aspx
and/or consult your colleagues and supervisor regarding developmental stages, how to rate clinical cases or how to use this instrument.

Therefore, underpinning the use of HoNOSI is considering the age and developmental stage of the infant or young child and the typical parameters of child development. For pre-term infants, the clinician is likely to consider whether developmental age is more relevant than birth age. Certain behaviours e.g. oppositional or disruptive behaviours are normal in the development of young children as they explore and begin to individuate, as are separation anxiety symptoms late in the first year. The capacity to learn to regulate their emotions and behaviour occurs within the context of their primary caregiving relationships.

The HoNOSI has been developed for use by clinicians working both in specialist perinatal and infant mental health units and for those who see infants and young children in child and adolescent mental health service settings. It is expected that knowledge of development and the impact of the early years on later development is central to all clinicians working with infants, pre-schoolers, children and adolescents although specialist infant mental health and generalist child and adolescent mental health service staff may have differing levels of familiarity and specialist knowledge. While the prime audience are clinicians within the specialist mental health sector (both infant and CAMHS), it is possible that HoNOSI may also be of use to clinicians in the primary mental health sector.

For clinicians, the HoNOSI aims to:

- inform treatment decisions by highlighting unexpected progress or deterioration for infants/young children;
- document the progress of the infant/young child and make overt changes, and;
- facilitate discussion of infants and young children's presentation among clinicians.

Services might also use the HoNOSI as an adjunct to:

- describing program or service effectiveness, and;
- contributing to examining the local implementation of evidence based treatments.

As with HoNOSI's companion instrument, the Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA), rating the presence of difficulties on any HoNOSI scale does not necessarily indicate that the source of the difficulty, or the required intervention, is solely with the infant/young child. To interpret any clinical ratings as indicating the infant/young child is the sole locus of the psychopathology is to both misinterpret the instrument and misunderstand the interdependent nature of child development and the environment in which it occurs.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Background to HoNOSI

HoNOSCA was developed in the UK by a team led by Dr Simon Gowers and Prof Richard Harrington in the late 1990s. The aim was to have an instrument that was brief, reliable and valid enough that it could be used routinely in Child and Adolescent Mental Health Services (CAMHS) to ascertain both a snapshot of children and adolescent's level of symptomatology as well as change over time. The first HoNOSCA paper was published in 1999 and since then HoNOSCA has been used widely (Adamis et al., 2011; Bilenberg, 2003; Brann, 2010; Duffy & Skeldon, 2014; Gowers et al., 1999; Hanssen-Bauer et al., 2010; Kisely et al., 2010; Lesinskiene, Senina, & Ranceva, 2007). Australian CAMHS trialled HoNOSCA in 1999 and by 2003, it was part of the National Outcomes and Casemix Collection (NOCC) implemented by all Australian State and Territory Governments (Department of Health and Ageing, 2009).

The Australian Child and Adolescent Mental Health Information Development Expert Advisory Panel (CAMHIDEAP) recognised the need for related instruments for infants and young children under the age of 4 years as a clear gap in the NOCC protocol. The panel undertook a number of reviews of available outcome measures using both a rigorous process for assessing clinician and parent rated instruments and parameters for decision making. A significant gap in outcome measurement was identified. One option that ran as an ongoing subtext was to create a HoNOSCA like instrument for the under 4 year olds to fill this gap. This was not intended to replace any other instruments but rather to fill the gap in clinician-based instruments that exist in the public domain, and could readily be incorporated into the existing extensive collection and reporting infrastructure across the multiple governments and organisations that comprise the Australian mental health system.

Given the central importance of life-span approaches to mental health practice and service delivery and the emerging significance of the impact of the life cycle in mental health, addressing this gap is a national priority. With a collaborative relationship developed during an international examination of HoNOSCA (Hanssen-Bauer et al., 2007), the lead initiative occurred from Dr Sally Merry of New Zealand, with in principle support from key figures of the collaboration from the UK, Denmark, Norway as well as Australia. With the initial ideas scoped by Dr Merry, the CAMHIDEAP facilitated further development of HoNOSI. The working group included Dr Peter Brann, Dr Elisabeth Hoehn, Ms Margaret Hoyland, Dr Anne Sved Williams, Dr Nick Kowalenko, Ms Wendy Preston, Mr Tim Coombs and Ms Rosemary Dickson.

This version was developed to have a starting point for empirical testing and anticipated modification and parallels the HoNOSCA in so far as the quality of the ratings is more obviously a function of the knowledge of the rater and makes overt the rater's understanding of the infant or child's presentation. As with HoNOSCA, presentations do not differentiate aetiology. For example, an infant's apparent anxiety may be thought to be a reflection of the parent's anxiety but the anxiety will still be rated. No rating should ever be understood to comment on the cause or the solution. A clinical rating is simply shorthand for indicating there are difficulties: how best or who is best to deal with them, is a separate question.

For ease of reading, unless a specific term is required, 'infant' is used to refer to all children aged 0-47 months in this glossary. Unless specifically noted, parent should be read as including all caregivers. The term 'parents' may indicate plural or singular.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

The HoNOSI

Scale 1	Problems with disruptive behaviour/irritability/under controlled emotional regulation
Scale 2	Problems with activity levels, joint and/or sustained attention
Scale 3	Non accidental self-injury or lack of self-protective behaviours
Scale 4	Problems with feeding and eating behaviour
Scale 5	Problems with developmental delays
Scale 6	Problems with physical illness or disability
Scale 7	Problems associated with regulation and integration of sensory processing
Scale 8	Problems associated with sleep
Scale 9	Problems with emotional and related symptoms or over-controlled emotional regulation
Scale 10	Problems with social reciprocity
Scale 11	Problems with age appropriate self-care and environmental exploration
Scale 12	Problems with family life and relationships
Scale 13	Problems with attending care, education and socialisation settings
Scale 14	Problems with knowledge or understanding about the nature of the infant's difficulties
Scale 15	Problems with lack of information, understanding about services, or managing the infant's difficulties

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Key principles for rating the HoNOSI

- The rating period is the previous two weeks.
- Rate each scale in order from 1 to 15.
- Use all available information in making your rating. That is, HoNOSI ratings should reflect your judgement based on all sources of information available to you.
- Underlying problems may manifest across different scales but do not include a manifestation already rated in an earlier scale.
- Rate the most severe occurrence of the problem in the rating period, not the inferred cause.
- Clinically significant symptoms should be rated at a 2 or above.
- Ratings are informed by familiarity with, and a good understanding of, infant and child development. Good clinical practice recognises the role of supervision, team reviews, training and other resources as tools for reflecting on infant and child development. HoNOSI will not replace these. All examples in the glossary should be seen in a developmental context.
- The instrument is focussed on the infant but any rating may also reflect the relationship with the parent(s). Include the infant's temperament, parent's responses, interactions and level of distress regarding aspects of the infant's behaviour and presentation. As with HoNOSCA, the presence of a clinically significant rating (scoring 2, 3 or 4) does not imply that the source of the problem and/or the locus of intervention, is necessarily exclusively with the infant.
- When establishing a rating point, it can be useful to consider the underlying construct of a problem as a continuum.
- Differences between rating points may be influenced by the intensity of a problem, the presence of multiple problems, or the frequency of the problems.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

- Each item is rated on a five-point scale of severity (0 to 4) as follows:

Score	**Rating**	***Suggested implications***
0	No problem.	In the clinician's considered opinion.
1	Minor problem requiring no formal action.	Sub-clinical problem, evidence of some behavioural disturbance or distress, unlikely to be monitored or included in care plan.
2	Mild problem.	Mild problem, clinically significant issue, evidence of distress and/or behaviour disturbance. Likely to be monitored or included in care plan.
3	Problem of moderate severity.	Moderate problem, clinically significant issue, evidence of greater distress and/or, behavioural impact. Definitely monitored and included in care plan.
4	Severe to very severe problem.	Severe clinical problem, distress and/or behavioural disturbance dominant aspect of presentation, greater frequency and/or intensity of clinical activity as evidenced in care plan.
9	Not known or not applicable.	

- Higher ratings can be expected to accompany more severe, more frequent and more widespread presentations.

- As far as possible, the use of rating point 9 should be avoided, because missing data makes scores less comparable over time or between settings.

Specific information on how to rate each point on each item is provided in the Glossary.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 1: Problems with disruptive behaviour/irritability/under controlled emotional regulation	
Include	Oppositional and disruptive behaviours are normal in the development of young children as they explore and begin to individuate. This scale addresses problems with the age and developmentally appropriate capacity of the infant to manage strong feelings, without recourse to age inappropriate levels of overt disruptive behaviours. Clinically, the identification of age-inappropriate emotional regulation does not indicate the source of any difficulties. It may be expected, though not invariable, that regulating emotions connected to hunger, tiredness, and separation may be more prominent with younger infants while overt aggression or rage may be prominent with the older children. Include behaviour associated with any disorder (such as hyperkinetic disorder, depression, autism). Include the capacity to manage intense feelings of hunger, tiredness or separation from the primary caregiver. Include difficulty calming, demanding, whining, undue irritability, excessive crying, frequently arching back and stiffening coupled with turning away from all eye contact, physiological indicators of stress (hiccups, yawns, non-injurious scratching) and manifestations of under controlled emotional regulation. Include physical or verbal aggression (e.g. pushing, hitting, biting, kicking, teasing), to others (e.g. children, parents or other caregivers, siblings, familiar adults or strangers),
Exclude	Problems associated with feeding and sleeping rated at scale 4 (feeding) and scale 8 (sleeping). Problems directly associated with physical health illnesses or disability rated at scale 6. Problems associated with self-injury rated at scale 3. Problems associated with over-controlled emotional regulation or inhibited behaviours are rated at Scale 9. Problems associated with Sensory Processing that impact on adaptive interaction are rated at Scale 7.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action.
2	Mild problem. May be limited to one context.
3	Problem of moderate severity with disruptive or aggressive behaviour or under controlled emotional regulation. May be/likely to be in more than one context.
4	Severe to very severe problem with disruptive or aggressive behaviour or under controlled emotional regulation. May occur in almost all activities.

Scale 2: Problems with activity levels, joint and/or sustained attention	
Include	Include problems with overactivity/underactivity, joint and sustained attention associated with any cause, including related to aspects of the caregiving environment (e.g. lack of appropriate stimulation, opportunities for motor development). Include problems with restlessness, fidgeting, jerkiness, distractibility, listlessness or concentration due to any cause, including depression. Include issues of sustained as well as joint attention. Activity and attention difficulties may manifest in altered levels of vigilance, impaired turn taking in behavioural interactions, pronounced startle reflexes and rigidity. Where two factors appear to negate each other (e.g. joint attention problematic but sustained attention is not problematic), rate the most severe occurrence.
Exclude	Problems directly associated with physical health illnesses or disability scored at scale 6.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action.
2	Mild problem with overactivity/underactivity or restlessness but with age-appropriate support/structure, the infant can modify their activity levels. Some vulnerability in joint and/or sustained attention however the infant's development is only mildly affected.
3	Problem of moderate severity with overactivity/underactivity. Activity levels may be difficult to control even with appropriate supports. May be significant issues with joint and/or sustained attention.
4	Severe to very severe problem with overactivity/underactivity. Likely to be impacting negatively on the infant's capacity to engage and achieve developmental milestones across multiple contexts. Consistent and severe limitations in joint and/or sustained attention.

Scale 3: Non accidental self-injury or lack of self-protective behaviours	
Include	With infants and pre-schoolers, the question of intentionality is less clear than with older children. While intention should be considered, it will not always be apparent and the clinician may draw on clinical experience to infer intentionality. Behaviours included here are essentially those that result in self-harm that are not the consequence of an accident. However, self-injurious behaviours and actions are rated here irrespective of any indication of intent. May include self-soothing behaviour that results in injury or harm e.g. hitting, biting, hair pulling, head banging, rocking, cutting, scratching, excessive sucking leaving marks, skin scratching or picking. May include lack of self-protective reflexes, inhibition of pain and reassurance responses e.g. when an infant is clearly hurt yet inhibits a response where other infants of the same age would be expected to cry, flinch and look to parent(s) for reassurance. Include attempts to stab self with a pen or other non-lethal object, cutting self with knives or scissors, deliberately jumping from a height with injurious intent, frequently discussing intent to self-injure. May include consideration of behaviours during play.
Exclude	Self-injurious behaviour secondary to a medical condition. Accidental self-injury unless clearly from a lack of self-protective reflexes.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action with lack of self-protective reflexes or self-injurious behaviours.
2	Mild problem with self-injury or a lack of self-protective reflexes. May include rubbing, scratching, rocking or play which leads to mild levels of physical injury. Play that regularly involves self-injury. Occasional episodes where self-protective reflex is inhibited.
3	Problem of moderate severity with potential or actual self-injury. May include moderately severe problems with a lack of self-protective behaviours that lead to, or potentially lead to injury. May be preoccupation, repeated episodes, or inhibition of pain responses to self-injury.
4	Severe to very severe self-injury occurs. Episodes of physical self-injury. May include inhibition of response to pain/discomfort and lack of self-protection and self-soothing leading to severe self-injury.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 4: Problems with feeding and eating behaviour	
Include	Feeding behaviours progress with development. The acknowledgement of problems in this area will be influenced by the duration, distress and incongruence of the concerning behaviours with the infant's age and age appropriate development. Include problems related to difficulties with breast feeding, bottle feeding and solids. Include all feeding difficulties irrespective of potential cause or solution. Nutritional difficulties may not always be present but should be considered. Include behaviours such as reluctance, resistance or refusing to feed; tiring or sleeping readily during feeding; feeding related distress (e.g. fussiness or crying); maintaining adequate nutrition which may result in nasogastric/gastrostomy tube feedings; sensory adversity; vomiting and difficulty in achieving developmentally appropriate food or feeding skills e.g. limited diet, consistent refusal of certain foods, groups, or types (e.g. solids), or modes of eating e.g. refusal to eat independently; little recognition of the relationship between hunger, feeding and satiety. Include under- and over-eating. Include feeding problems related to prematurity, physiological problems and gastrointestinal symptoms.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action. Problems may be transient and may be expected at the infant's developmental stage.
2	Mild problem with feeding or eating. Nutritional intake is likely to be within expected parameters.
3	Problem of moderate severity with feeding or eating. Some risk of nutritional problems.
4	Severe to very severe problem with feeding or eating.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 5: Problems with developmental delays.	
Include	Include problems with developmental delays not rated at other scales. Delays may occur in areas such as cognitive, motor, language, or communication development. Concerns should be rated both irrespective of cause and whether additional professional assessment or intervention has occurred (e.g. paediatrics, speech pathology). It may be difficult to distinguish one domain from another. Cognitive, motor and communication difficulties may manifest in balance, coordination, proprioception, problem solving, articulation, comprehension, sentence structure, vocabulary, communication pragmatics, gestures, vocal quality or range, interference with vocalisation (e.g. dummy, fingers). Difficulties in these areas may impact on ability to interact effectively with the environment and themselves in the areas of communication, motor and cognitive skills. While corrected age is a useful construct with premature infants, chronological age may be the more useful in identifying potential need for intervention.
Exclude	Physical illness or disability problems such as vision and hearing problems (rated at scale 6).

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action. These may be expected to be within the typical range of development.
2	Mild problem that may be noted across more than one setting and in comparison to similar aged peers.
3	Problem of moderate severity that may be noted across settings compared with similar aged peers.
4	Severe to very severe problem with cognitive, motor or communication skills. Likely to cause significant distress for the infant and/or family. May be severe delays compared to similar aged peers.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 6: Problems with physical illness or disability	
Include	Physical health problem or disability which limits or prevents movement, impairs sight or hearing or otherwise interferes with functioning. Problems in this area may be observed or based on reports from others. Include side effects from medication, physical effects from drug/alcohol exposure, or physical complications of psychological disorders. Include physical complications or disability as consequence of self-injury. Ratings will be influenced by consideration of impact of illness on everyday functioning.
Exclude	Problems with cognitive, motor or communication skills already rated at scale 5.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action (e.g. cold, non-serious fall, teething). Parent voices concern about transient physical illness or physical symptoms but these are not considered serious by the parent or clinician.
2	Mild problem with physical illness or disability, which may occasionally prevent or challenge engagement in usual activities. Overall structure of their day is typically preserved and activities such as the ability to play are only mildly affected.
3	Problem of moderate severity with physical illness or disability, resulting in some ongoing distress and loss of function. Typically, there is some time each day, in which they are able to engage in usual activities, such as play.
4	Severe to very severe problem with physical illness or disability that result in serious distress and/or loss of function. Normal everyday routines and activities, including play, are seriously impacted because of the physical problem. Considerable input of effort and resources may be required to care for the infant and support the parent.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 7: Problems associated with regulation and integration of sensory processing	
Include	Problems associated with processing, regulating and integrating information from sensory stimuli which interfere with the sensory regulation required for adaptive interaction with and exploration of the world. While problems with sensory organs are rated at scale 6, this scale is more concerned with the processing of otherwise apparently intact sensory organs. Problems associated with sensory processing can reflect hypersensitivity (over-reactive therefore avoidant or fearful/cautious) and/or hyposensitivity (under reactive therefore seeking or impulsive) to one or more normal sensory stimuli. Sensory stimuli include vision, touch, hearing, taste, smell and spatial awareness including the sensation of movement and awareness of body position in space. Problems associated with the regulation and integration of sensory processing usually occur across multiple settings and within multiple relationships. Intensity, frequency, duration and location of problematic sensory stimuli may impact on the infant's presentation. Examples of the manifestation of sensory regulation difficulties may include responsiveness to fabrics, movement, travel, focus on apparently irrelevant objects and an avoidance of play. They may appear to have a preference for swaddling, or to seeking or avoiding certain fabrics.
Exclude	Problems with disruptive behaviour/under controlled emotional regulation rated at scale 1. Problems with activity and attention levels rated at scale 2. Problems with feeding rated at scale 4. Problems associated with cognitive, motor or communication difficulties rated at scale 5. Problems with physical illness or disability rated at scale 6. Problems with anxiety and depression and over controlled emotional regulation rated at scale 9.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action with sensory processing (over or under responding to normal sensory stimuli). However, the impact on adaptive daily functioning and exploration of the world is typically minor.
2	Mild problem with sensory processing identified and impacting on the infant. The infant and/or family may be showing signs of distress but maintaining appropriate developmental milestones. Definite and minor impact on functioning in daily tasks or in maintaining interactions in primary care-giving relationships. May become agitated, distressed, or disengaged when exposed to specific sensory stimuli.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

3	Problem of moderate severity with sensory processing that are impacting on the infant's capacity to engage with their environment. May manifest as diminished exploration and play. Expect a definite and moderate impact on daily functioning.
4	Severe to very severe problem related to sensory processing directly impacting the infant's social, emotional and physical wellbeing. Definite and severe impact that is typically ongoing.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 8: Problems associated with sleep	
Include	Sleep disturbance is common for infants. Include difficulties in both settling and maintaining sleep irrespective of where the locus of the difficulty is thought to be (infant, parent, living arrangements). Include excessive sleep (e.g. which interferes with opportunities for skills or social development), insufficient sleep (e.g. periods of awakenings or reduced sleep time), disturbed sleep (e.g. sleep talking, sleep walking, night terrors, or any other disturbance during sleep when the infant does not seem to respond to the parents) or difficulties resettling. Include snoring or loud mouth breathing with breath holding or gasping.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action. Typically within expected developmental norms, infrequent and where the family appear to have some approaches that successfully address the problem.
2	Mild problem which is intermittent. The family appear to have some success in addressing the problem for the infant.
3	Problem of moderate severity. The infant‘s sleeping pattern is of concern to the parents or family or it is likely to be interfering with functioning or development. The sleep disturbance occurs frequently and may be significantly out of keeping with age expectations.
4	Severe to very severe problem. The sleeping pattern is a cause for great distress in the parents and family and may be significantly out of keeping with age norms. The sleep disturbance is present nearly all the time and significantly interferes with functioning or development.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 9: Problems with emotional and related symptoms or over-controlled emotional regulation	
Include	Symptoms of depression, anxiety and phobias. Problems with negative or inhibited affect in the infant suggestive of low mood, anxiety, fear, emotional withdrawal, or over-controlled emotional regulation. May include fears, anxiety or emotional withdrawal from parents and others. Include incongruent lack of emotional expression. May be expressed with changes in curiosity, clinging, masking face, incongruent emotional expression, crying, anger, hypnotic gaze, withdrawal and blank expression, exaggerated positive or negative emotional responses. May include excessive stillness, frozen watchfulness, quiet rage and restrictions in affect range. An apparent increased tolerance for aversive adult behaviour, or problems seeking appropriate comfort or safety should be considered. Include age or developmentally inappropriate lack of wariness, or avoidance of parents.
Exclude	Physical sequelae of psychological disorders or medication – rated at scale 6. Disruptive behaviours resulting from emotional distress – rated at scale 1. The emotion associated with the disruptive behaviour is rated here at scale 9.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action, or transient mood, anxiety and emotional symptoms or changes.
2	Mild problem with emotional symptoms.
3	Problem of moderate severity with emotional symptoms which are preoccupying, intrude into some activities and are uncontrollable at least sometimes.
4	Severe to very severe problems with emotional symptoms which intrude into all activities and may be nearly always uncontrollable.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 10: Problems with social reciprocity	
Include	This scale addresses the infant's engagement in, and engagement of others in, age and developmentally appropriate interactions. There may be problems with seeking, engaging and enjoying interactions with familiar adults and children, including development of the social smile at 6 weeks. Responses to social engagement or social intrusion from others may not be responded to appropriately e.g. ambiguous half smiles. Problems may manifest with reciprocity in communication, play, and games. Reciprocity may be expressed both pre-verbally and verbally, as well as behaviourally. Problems may manifest as indiscriminate and overfamiliar social interactions as well as withdrawn and disengaged social interactions. Problems rated in this scale may include the infant's capacity to manage appropriate eye contact e.g. the infant may not gaze at the parent's face or at an interesting object when shown. Problems may include not socially referencing others, brief glances without sustained looking (difficulty gaining and sustaining eye contact); avoidant gaze; no eye contact (but no active avoidance either) and unfocused eyes. Problems with vocalisations relating to reciprocity of interactions, such as turn taking, engagement attempts, and vocal mirroring may also be relevant indicators of social reciprocity issues.
Exclude	Difficulties with communication separate to the social reciprocity function are rated at Scale 5. Difficulties with the emotional attunement of parent's and carers to the infant and misalignment between the infant's needs and the parents' or carers' responses should be rated at Scale 12.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action. Transient or mild problems in the infant's developing capacity to engage in social relationships.
2	Mild problem.
3	Problem of moderate severity with social reciprocity.
4	Severe to very severe issues with social reciprocity. Problems likely to occur in many areas, over time and intrude across most interactions.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 11: Problems with age appropriate self-care and environmental exploration	
Include	This scale addresses age-appropriate self-care and exploration of the environment. Self-care is more likely to be a prominent consideration with older children. Self-care is likely to include age appropriate levels of assistance with bathing, feeding, dressing, playing etc. Problems with self-care and environmental exploration may exist due to environmental restrictions, including parent's comfort, concerns or control. Include problems with activities of self-care such as washing, dressing, toileting. Exploration may include visual, tactile, verbal as well as physical exploration (under or over exploration). Include problems with complex skills such as play, autonomous activities or separating from parents, taking into account the norm for the infant's age and developmental stage. Difficulties may be indicated by regression to an earlier stage of development. The impact on exploration and self-care resulting from separation problems with parents when the infant is attending structured socialisation settings (e.g. day care, pre-school) may be rated here although actual attendance issues are rated at scale 13. Include poor levels of functioning arising from apparent lack of motivation, mood, environmental restriction or any other issue whether it is considered to arise from the infant, parents or the environment.
Exclude	Do not include feeding problems rated at scale 4. Do not include sleeping problems rated at scale 8. Do not include lack of opportunities for exercising intact abilities and skills, as might occur in an over-restrictive family rated at scale 12. Do not include the outcome of limited environmental exploration on structured socialisation settings rated at scale 13.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action with self-care or exploration of the environment.
2	Mild problem with self-care or exploration of the environment.
3	Problem of moderate severity with self-care or exploration of the environment.
4	Severe to very severe problem with self-care or exploration of the environment that is likely to be intruding across settings, activities and persons.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 12: Problems with family life and relationships	
Include	This scale addresses problems in family life that are thought to impact on the infant. If the parents are separated, consider the relationship with each parent and the separated parents' ability to co-parent where appropriate. Include relationships with significant others – grandparents, siblings, extended family members, child care providers. Include instances of neglect including physical (e.g. lack of sufficient access to appropriate food, shelter and clothing) and emotional (e.g. lack of warmth, comfort and age appropriate regulation of the infant's affect). Parental reflective capacity; the availability of access to caring, attentive and empathic adults and the ability to keep the infant in mind, should be considered. Include parent or family irritability with the infant. Difficulties in managing powerful emotions or any consequent harmful behaviour by those in the infant's immediate environment should be considered. Include instances of physical or verbal hostility or abuse towards the infant, as well as family hostility or conflict which impacts on the infant. Consider capacity for significant others to contain powerful negative emotions towards the infant. Issues such as parental or sibling mental health, substance use and personality problems should be included if they have an effect on the infant.
Exclude	Do not include disruptive behaviour by infant, rated at scale 1. Do not include problems with social reciprocity rated at scale 10.

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action. Some concerns about family relationships may be evident but effect on infant mitigated through adequate parental reflective capacity and action are apparent.
2	Mild problem with family relationships. Some impact on the infant's development is apparent.
3	Problem of moderate severity with family relationships. Considerable impact on infant development apparent.
4	Severe to very severe problem in family relationships with severe impact on the infant.

Scale 13: Problems with attending care, education and socialisation settings	
Include	This scale addresses attendance at the prime socialisation setting outside of the immediate family. Include attendance at any type of regular socialisation and care activity at the time of rating e.g. regular care with extended family or formal early childhood education (sometimes called kindergarten or pre-school). Include activities irrespective of location or whether a family member is present (e.g. regular play group sessions at infant's home). Include refusal of, or withdrawal from early childhood education, childcare, play group or similar regular socialisation activity, irrespective of cause. Include limited or minimal opportunities to attend socialisation activities appropriate to the infant's age. Include consideration of additional supports such as reassurance, transitional objects, required to settle the infant in the setting. If early childhood education, childcare etc. is in holiday break, rate the last two weeks of the previous open period. Note: Infants and young children will communicate their reluctance and distress at attending these settings through a range of symptoms. These may include problems in feeding, toileting, eating, playing, communicating and sleeping both at the settings and around the transition time. These symptoms in themselves are likely to be rated at different HoNOSI scales and are not the sole source of rating at this scale. However, it is acknowledged that the reluctance to attend may be conveyed to the clinician through these symptoms. The actual attendance problems are rated at this scale. Note: Reluctance to attend a socialisation setting may reflect problems in that setting for the infant. Reluctance to attend may also occur in the dyadic relationship or simply from parental concerns. To reiterate, acknowledging a problem does not mean that the source of the problem or the required solution necessarily lies with the infant. HoNOSI is agnostic as to the locus of any intervention.
Exclude	Many infants have not attended a socialisation setting outside the family and the clinician will typically not consider this a problem. However, a clinician may decide to rate this non-attendance as a problem; for example, where non-attendance has been considered to reflect extended separation difficulties. All behaviours and emotional expressions or consequences of problems associated with attendance or separation are rated at their respective scales (e.g. Disruptive at scale 1, Feeding at scale 4, Emotional at scale 9, Environmental exploration at scale 11). Absences due to illness of infant or parents requiring them to be absent from the setting. This typically includes medical conditions, such as fevers, contagious illnesses or infections which would be rated at scale 6.

Rating	**Description**
0	No problem. Infant displays age appropriate behaviour on separation from their parents and settles readily when comforted in the environment.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

1	Minor problem requiring no formal action with attending and may display reluctance for brief periods. Responds with small amount of support additional to that typically required at this age.
2	Mild problem with some sessions missed or refusal to participate in activities when attending.
3	Problem of moderate severity with several days missed during rating period due to infant's reluctance to attend.
4	Severe to very severe problem with infant absent for most of the days or sessions during the rating period.

Appendix 1: The Health of the Nation Outcome Scales for Infants (HoNOSI)

Scale 14: Problems with knowledge or understanding about the nature of the infant's difficulties	
Include	Include lack of useful information or understanding available to the parents, caregivers, referrers or support system about the nature of the difficulties. Include problems with capacity or knowledge to understand the infant's difficulties. Include limited or incorrect understanding of the infant's developmental stage and needs. Include misunderstanding, minimising, elaborating or exaggerating the difficulties, impact or distress as well as inaccurate attribution of the infant's difficulties. Include lack of explanation about the difficulty/diagnosis, the cause of the problem or understanding of the prognosis or the impact on the infant. Rating a problem here does not exclude the service system revising their understanding of the infant's difficulties. In many ways, problems rated here may indicate a lack of congruence between the parent's and other's views about the nature of the difficulties and the views of the clinician (or the assessing or treating system.)

Rating	**Description**
0	No problem. Parents, referrers or carers demonstrate a good level of understanding about the difficulties.
1	Minor problem requiring no formal action. For example, parents essentially understand infant's difficulties but with occasional misunderstandings such as sometimes downplaying, or exaggerating the infant's difficulty or distress.
2	Mild problem in understanding infant's difficulties.
3	Problem of moderate severity. Parents have very little or very poor knowledge about the nature of their infants' problems.
4	Severe to very severe problem. For example, parents have no understanding about the nature of their infant's problems. Significant disagreement between the parents, or the referrer's or the carer's views and the views of the assessing or treating system.

Scale 15: Problems with lack of information, understanding about services, or managing the infant's difficulties	
Include	Include lack of useful information available to the parents, caregivers, or referrers, or a lack of understanding regarding services or management of the difficulties. Include parental willingness or ability to utilise services or interventions to support the infant. The consistency with which parent's understand or use appropriate strategies and the extent to which supports are required to help the parent's use optimal approaches may be considered. Include parents, referrers or carers use and implementation of information and appropriate and feasible strategies. Include problems with accessing available services appropriate to the infant's difficulties (e.g. early childhood nursing, child protection, family support). Rating a problem here does not exclude the service system revising their understanding of the optimal approach to managing the infant's difficulties. In many ways, problems rated here may indicate a lack of congruence between the family, carer's or referrer's views about the management of the infant's difficulties and the views of the clinician (or the assessing or treating system's views).

Rating	**Description**
0	No problem.
1	Minor problem requiring no formal action. For example, parents have an adequate understanding of how best they and other resources can help their infant, or they are actively seeking appropriate information, support or access to services.
2	Mild problem in understanding or willingness to use the appropriate services, approaches, resources and supports for the infant's difficulties.
3	Problem of moderate severity in understanding or willingness to use the appropriate services, approaches, resources and supports for the infant's difficulties.
4	Severe to very severe problem in understanding or willingness to use the appropriate services, approaches, resources and supports for the infant's difficulties.

Appendix 2: Parent-Infant Relationship Global Assessment Scale (PIR-GAS)

PIR-GAS (Parent-Infant Relationship Global Assessment Scale)

The PIR-GAS allows for a judgement about the infant-parent relationship. On the PIR-GAS scale, the quality of the infant-parent relationship, ranges from well adapted to severely impaired. The clinician typically completes the scale after multiple clinical evaluations for a referral problem. Clinicians who use the PIR-GAS should remember that relationship problems may or may not co-occur with symptomatic behaviours in the infant. In other words an infant may have symptoms of a serious mental health disorder and yet have adaptive, flexible relationships with parents and important other adults. Diagnosis of relationship disturbances or disorders are made not only on the basis of observed behaviour but also on the basis of parent's subjective experience of the child as expressed during a clinical interview and the subjective experience of the child, as expressed in a play interview, for example. When difficulties in the focal relationship are apparent, the clinician assess the intensity, frequency and duration of the difficulties in order to classify the relationship problem as a perturbation, a disturbance, or a disorder.

PIR-GAS scores are classified as:

81-100 – Adapted Relationship
41-80 - Features of a Disordered Relationship
0-41 - Disordered Relationship

A PIR-GAS score below 40 indicates a relationship disorder. However, many infant-parent relationships with a PIR-GAS score between 40 and 80 may show tendencies toward, or features of, a disordered relationship that may benefit from therapeutic intervention.

It is not necessary to know the etiology of current relationship problems in order to use the scale. Symptoms may derive from conditions within the infant, from within the caregiver, from the unique 'fit' between infant and caregiver, from the larger social context, or from a combination of several on these factors. The clinician who understands the stressors that are affecting an infant-parent relationship may know a great deal about the origins of the problem. Moreover, coding of the infant-parent relationship on the PIR-GAS does not imply that the current nature and quality of the infant-parent relationship is immutable. PIR-GAS scores are meant to capture the nature of the relationship at time of assessment. Scores may vary over time, as the quality of a relationship is subject to numerous factors, both intrinsic and extrinsic to the relationship, including therapeutic intervention. Additionally, a clinician should be aware of how particular developmental stages in the child may adversely interact with a parent's experiences, expectations, or challenges to produce disturbances in the infant-parent relationship. These challenges may evolve or go into remission as the infant (or parent) changes.

The clinician should use the PIR-GAS to assess the relationship between primary caregiver(s) and the infant. Primary caregivers may be biological, foster, and adoptive parent(s), as well as grandparents, members of the extended family, and caregivers outside the family.

In assessing the parent-infant relationship, the clinician should consider multiple aspects of the relationship dynamic, including:

1. Overall functional level of both the child and the parent
2. level of distress in both the child and the parent
3. adaptive flexibility of both the child and the parent
4. level of conflict and resolution between the child and the parent
5. effect of the quality of the relationship on the child's developmental progress

DC:0–3R

Parent-Infant Relationship Global Assessment Scale (PIR-GAS)

91–100 Well Adapted Parent–child relationships in this range are functioning exceptionally well. They are mutually enjoyable and without sustained distress. They evidence adaptation to new circumstances and are typically free of conflict as parent and child manage the stresses of everyday life. The relationship clearly promotes the growth of both partners.

81–90 Adapted Relationships in this range are also functioning well, without evidence that the relationship is significantly stressful for either partner. Interactions within these relationships are frequently reciprocal and synchronous, without distress, and reasonably adaptive. At times parent and child may be in substantial conflict, but conflicts do not persist longer than a few days and are resolved with appropriate consideration of the child's developmental status. The pattern of the relationship protects and promotes the developmental progress of both partners.

71–80 Perturbed Some aspect of the overall functioning of relationships in this range is less than optimal; partners may experience transient distress lasting up to a few weeks. Nevertheless, the relationship remains characterized by adaptive flexibility. The disturbance is limited to one domain of functioning. Overall, the relationship still functions reasonably well and does not impede developmental progress.

Example: An infant with a minor physical illness sleeps poorly for several nights, exhausting his parents; or parents moving into a new house are less attentive to their infant, who is less able to self-regulate in the unfamiliar new surroundings.

61–70 Significantly Perturbed Relationships in this range of functioning are strained but still largely adequate and satisfying to the partners. Conflicts are limited to one or two problematic areas. Partners may experience distress and difficulty for a month or more. The relationship maintains adaptive flexibility, as parent and child seem likely to negotiate the challenge to their relationship successfully. A parent may be stressed by the perturbation, but is not generally overconcerned about the changed relationship pattern, considering it within the range of expectable, relatively short-lived difficult periods in a lifelong relationship.

Example: Following the birth of a new sibling, a toddler develops new-onset food refusal and a sleep disturbance that lasts more than a month.

51–60 Distressed Relationships in this range of functioning are more than transiently affected as one or both partners experience distress in the context of their relationship. Parent and child maintain some flexibility and adaptive qualities, but conflict may spread across multiple domains of functioning, and resolution is difficult. The developmental progress of the dyad seems likely to falter if the pattern does not improve. Caregivers may or may not be concerned about the disturbed relationship pattern. Neither parent nor child is likely to show overt symptoms resulting from the disturbance.

Example: A child expresses distress and oppositionality during toilet training and feeding. Her mother is increasingly worried about her ability to engage her daughter in these activities in growth-promoting ways.

continued

7

DC:0–3R

Parent–Infant Relationship Global Assessment Scale (PIRGAS)

41–50 Disturbed The adaptive qualities of a disturbed relationship are beginning to be overshadowed by problematic features. Although not deeply entrenched, dysfunctional patterns appear more than transient. Developmental progress can still proceed, but may be temporarily interrupted.

Example: A parent and child engage in excessive teasing and power struggles in during, feeding, dressing, and bedtime. Although parent and child attempt pleasurable interactions, their teasing often goes too far, leaving one or both partners distressed.

31–40 Disordered Rigidly maladaptive interactions, particularly if they involve distress in one or both partners, are the hallmark of disordered relationships. Most interactions between partners are conflicted; some relationships without overt conflicts may nevertheless be grossly inappropriate developmentally. Developmental progress of the child and the parent–child relationship is likely to be influenced adversely.

Example: A depressed parent repeatedly seeks comfort from her infant, actively recruiting caregiving behavior from the child. The child's engagement in exploratory play is limited.

21–30 Severely Disordered Relationships in this range of functioning are severely compromised. Both partners are significantly distressed by the relationship itself. Maladaptive interactive patterns are rigidly entrenched. To an observer, interactive patterns seem to have been in place for a long time, although the onset may have been insidious. In a severely disordered relationship, a significant proportion of interactions are likely to be conflicted. Developmental progress of the child and the relationship is clearly influenced adversely. Indeed, the child may lose previously acquired developmental skills.

Example: A father and his toddler frequently interact in a conflicted manner. The father sets no limits until he becomes enraged. Then he spanks the toddler vigorously. The toddler is provocative, and the father feels angry with him all the time.

11–20 Grossly Impaired Relationships in this range of functioning are dangerously disorganized. Interactions are disturbed so frequently that the infant is in imminent danger of physical harm.

1–10 Documented Maltreatment The relationship contains documented neglect and physical or sexual abuse that is adversely affecting the child's physical and emotional development.

Reprinted from
Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood, Revised (DC:0-3R), by ZERO TO THREE (2005)
Order at www.zerotothree.org/bookstore.

Appendix 3: Clinical Worry and Severity Judgement Ratings

HoNOSI Field Trial Demographics and Ratings

Date: ____/_____/_____

Infant Demographics

1. Age (months):_____

2. Sex (please circle): M / F

3. Reason for referral:_____

Your Clinical Experience

4. How long have you been working with infants and pre-schoolers (please circle)?

<2 years 2-5years 5-10 years 10+ years

Your Profession

5. What is your profession (please circle)?

Psychiatrist Psychologist Nurse Social Worker OT Speech Pathologist Other

Clinical Concern Rating

6. Overall, how concerned are you about this infant?

0	1	2	3	4	5	6
Not concerned						Extremely concerned

Severity Judgement Rating

7. In your clinical judgement, how severe do you consider the infant's overall social and emotional problems?

0	1	2	3	4	5	6
No problem						Extremely severe problem

Appendix 4: Distribution of HoNOSI Ratings – Clinical Cases

Table A4.1: *Frequency Distribution of HoNOSI Severity Ratings for Clinical Cases (n = 108)*

	HoNOSI Scale	Severity Rating[1] 0 (%)	1 (%)	2 (%)	3 (%)	4 (%)	9 (%)
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	32.4	15.7	27.8	21.3	2.8	0.0
2	Problems with activity levels, joint and/or sustained attention	37.0	15.7	31.5	13.0	2.8	0.0
3	Non accidental self-injury or lack of self-protective behaviours	76.9	13.9	4.6	3.7	0.9	0.0
4	Problems with feeding and eating behaviour	41.7	24.1	20.4	12.0	1.9	0.0
5	Problems with developmental delays	52.8	15.7	15.7	13.0	2.8	0.0
6	Problems with physical illness or disability	71.3	18.5	8.3	1.9	0.0	0.0
7	Problems associated with regulation and integration of sensory processing	70.4	15.7	12	1.9	0.0	0.0
8	Problems associated with sleep	29.6	19.4	31.5	15.7	3.7	0.0
9	Problems with emotional and related symptoms or over-controlled emotional regulation	24.1	21.3	25.9	23.1	5.6	0.0
10	Problems with social reciprocity	42.6	18.5	21.3	15.7	1.9	0.0
11	Problems with age appropriate self-care and environmental exploration	58.3	18.5	14.8	8.3	0.0	0.0
12	Problems with family life and relationships	8.3	16.7	38.9	23.1	13.0	0.0
13	Problems with attending care, education and socialisation settings	55.6	17.6	19.4	5.6	1.9	0.0
14	Problems with knowledge or understanding about the nature of the infant's difficulties	22.2	16.7	27.8	22.2	11.1	0.0
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	29.6	27.8	15.7	23.1	3.7	0.0

[1]Severity Ratings:

0 - No problem
1 - Minor problem requiring no formal action
2 - Mild problem
3 - Problem of moderate severity
4 - Severe to very severe problem
9 - Not known or not applicable

Appendix 5: Spearman's Rho Correlation Matrix – Intercorrelation of HoNOSI Scales

The correlation matrix below (Table A5.1) shows that most correlations between the individual HoNOSI scales are statistically significant. The Grey shaded areas represent exceptions to these, which have a low correlation and are not statistically significant. Areas shaded in green show a correlation of above 0.5 and which are statistically significant (r=-0.73, p<0.01). The areas shaded in yellow show a correlation of r>0.6, p<0.01, however, these are HoNOSI total scores and their resulting statistic must be interpreted with caution as they are a sum of all HoNOSI scales 1 through 13 and as such, show a higher correlation than the individual scales, as excepted.

Table A5.1: *HoNOSI Scales Intercorrelation Matrix*

Spearman's rho Correlation Coefficient	Scale 1	Scale 2	Scale 3	Scale 4	Scale 5	Scale 6	Scale 7	Scale 8	Scale 9	Scale 10	Scale 11	Scale 12	Scale 13	Scale 14	Scale 15	HoNOSI Total
Scale 1: Problems with disruptive behaviour/irritability/under controlled emotional regulation	1.000															
Scale 2: Problems with activity levels, joint and/or sustained attention	.541**	1.000														
Scale 3: Non accidental self-injury or lack of self-protective behaviours	.413**	.453**	1.000													
Scale 4: Problems with feeding and eating behaviour	.303**	.275**	.118	1.000												
Scale 5: Problems with developmental delays	.440**	.483**	.213*	.272**	1.000											
Scale 6: Problems with physical illness or disability	.222*	.340**	.184	.116	.253**	1.000										
Scale 7: Problems associated with regulation and integration of sensory processing	.391**	.319**	.304**	.314**	.319**	.363**	1.000									
Scale 8:Problems associated with sleep	.489**	.527**	.229*	.513**	.277**	.249**	.270**	1.000								
Scale 9:Problems with emotional and related symptoms or over-controlled emotional regulation	.452**	.559**	.422**	.288**	.406**	.257**	.519**	.416**	1.000							
Scale 10: Problems with social reciprocity	.337**	.568**	.220*	.416**	.463**	.235*	.402**	.432**	.614**	1.000						
Scale 11: Problems with age appropriate self-care and environmental exploration	.359**	.464**	.216*	.193*	.305**	.266**	.435**	.155	.428**	.559**	1.000					

Appendix 5: Spearman's Rho Correlation Matrix – Intercorrelation of HoNOSI Scales

Scale 12: Problems with family life and relationships	.343**	.381**	.304**	.250**	.349**	.119	.324**	.331**	.514**	.416**	.207*	1.000			
Scale 13: Problems with attending care, education and socialisation settings	.319**	.401**	.288**	.116	.369**	.230*	.294**	.267**	.316**	.361**	.284**	.370**	1.000		
Scale 14: Problems with knowledge or understanding about the nature of the infant's difficulties	.470**	.573**	.341**	.338**	.387**	.235*	.368**	.455**	.471**	.496**	.427**	.611**	.481**	1.000	
Scale 15: Problems with lack of information, understanding about services, or managing the infant's difficulties	.382**	.551**	.323**	.215*	.240*	.199*	.404**	.413**	.520**	.436**	.361**	.510**	.388**	.706**	1.000
HoNOSI Total Scale 1-13	.694**	.769**	.468**	.534**	.641**	.399**	.593**	.651**	.771**	.743**	.567**	.586**	.545**	.690**	.601**	1.000

**. Correlation is significant at the 0.01 level (2-tailed)
*. Correlation is significant at the 0.05 level (2-tailed)

Appendix 6: Spearman's Rho Correlation Matrix – HoNOSI Scales with All Other Ratings

Table A6.1: *Spearman's Rank-Order Correlations: HoNOSI Scales and HoNOSI Total Score with the PIR-GAS, Clinical Worry and Severity Judgement Ratings (n = 108)*

HoNOSI Scale		PIR-GAS[1] r_s	Clinical Worry r_s	Severity Judgement r_s
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	-0.54	0.50	0.59
2	Problems with activity levels, joint and/or sustained attention	-0.57	0.57	0.66
3	Non accidental self-injury or lack of self-protective behaviours	-0.35	0.35	0.37
4	Problems with feeding and eating behaviour	-0.33	0.42	0.47
5	Problems with developmental delays	-0.52	0.57	0.57
6	Problems with physical illness or disability	-0.27	0.24	0.25
7	Problems associated with regulation and integration of sensory processing	-0.38	0.41	0.52
8	Problems associated with sleep	-0.48	0.51	0.56
9	Problems with emotional and related symptoms or over-controlled emotional regulation	-0.64	0.60	0.73
10	Problems with social reciprocity	-0.48	0.53	0.64
11	Problems with age appropriate self-care and environmental exploration	-0.31	0.33	0.36
12	Problems with family life and relationships	-0.58	0.71	0.60
13	Problems with attending care, education and socialisation settings	-0.36	0.40	0.45
14	Problems with knowledge or understanding about the nature of the infant's difficulties	-0.65	0.69	0.63
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	-0.57	0.63	0.65
Total Score	HoNOSI	-0.73	0.77	0.85

[1]There were 4 non-responses to the PIR-GAS

Appendix 7: Inter-rater Reliability Pilot Study

In order to test inter-rater reliability of the HoNOSI, multiple raters (clinicians) independently completed the HoNOSI, on the same 3 case vignettes. A copy of these vignettes are in Appendix 8.

Inter-rater reliability is defined as "the extent to which scores for patients who have not changed are the same for repeated measurement … by different persons on the same occasion" (Mokkink, Terwee, Patrick, et al., 2010, p. 743). Specifically, the inter-rater reliability of the HoNOSI was assessed by comparing independent clinical ratings for the same "individual consumer"; in this study, a set of case vignettes.

Method

Three case vignettes were written by clinicians from the project working group. As a preliminary test during the vignette development phase, members of the project working group then rated each of the case vignettes and compared their scores on each scale. Severity ratings that differed by more than 2 points on each individual scale were discussed amongst these test raters. Where necessary, appropriate amendments were made to the wording of the case vignettes. Appendix 8 shows the final versions of the three case vignettes used in the pilot study. Considering feasibility and other similar studies, the aim was to have 60-100 clinicians participate in rating case vignettes

Each clinician was given the set of these three vignettes. They were given a full copy of the HoNOSI to read and refer to whilst completing the vignettes. Clinicians were asked to independently rate each vignette using the provided HoNOSI and each vignette's HoNOSI rating sheet. Inter-rater reliability was assessed using both weighted kappa and intraclass correlations. For the latter, both absolute agreement and consistency of agreement were estimated. Findings are described in the *Results* section below.

Data Collection of Case Vignettes

In order to obtain a representative sample of clinicians across Australia to participate in the study, clinicians were approached from 5 jurisdictions; NSW, Qld, Vic, SA and WA (Table A7.1). Some jurisdictions had more than one service participating in the study (NSW from the outset, then later also Qld). In June 2017, a preliminary assessment on the number of completed sets of vignettes was made and it was determined that further data collection was needed in order to achieve a sample size with sufficient power to test levels of inter-rater reliability. At that time, ethics and site-specific approval was sought for a further two sites from Queensland to participate in the study. These additional sites provided 32% of the participating clinicians of 81 in total.

Data collection occurred over the period June 2017 through July 2018. As per ethics approval requirements, responses and the identity of individual participants was kept confidential. They were asked to complete the HoNOSI ratings independently of other study participants.

Table A7.1: *Number of clinicians participating in inter-rater reliability study per site*

Jurisdiction	Site	Clinicians n
NSW	Perinatal and Infant Mental Health/Child and Youth Mental Health Service, Northern Sydney LHD	7
NSW	Perinatal and Infant Mental Health Service, Western Sydney LHD	1
NSW	Mental Health Services, Western NSW LHD (NSW)	4
Vic	Alfred Health	12
Qld	Children's Health Queensland Hospital and Health Service	10
Qld	Townsville Hospital and Health Service (QLD)	11
Qld	Cairns Child and Youth Mental Health Service/Evolve	15
SA	Women's and Children's Health Network (CAMHS)	7
WA	Child and Adolescent Mental Health Services, Princess Margaret Hospital	14
Total		81

The final dataset consisted of 81 sets of 3 cases vignettes. Frequency distributions of HoNOSI severity ratings for each of the 15 HoNOSI Scales separately for each of the three vignettes are further detailed in Appendices 9 and 10. The entire range of clinical severity ratings, 0 through 4, was utilised on at least one of the three vignettes for each of the 15 HoNOSI Scales.

Less than 1% of all case vignette HoNOSI Scale ratings were rated as 9 *Not known/not applicable*. A breakdown of missing values is detailed in Appendix 11. Scales 11 *Problems with age appropriate self-care and environmental exploration* and 13 *Problems with attending care, education and socialisation settings* had the most missing values, though the number is very small (4 and 6 ratings respectively).

Vignette Descriptive Statistics

Table A7.2 and Figure A7.1 present summary descriptive statistics and the distribution of HoNOSI total scores for each of the 3 case vignettes. It can be seen in Table A7.2 that while there were some differences among the three vignettes in terms of overall clinical severity, the median HoNOSI total score was high, ranging from 26 to 32, notably higher than the median of 14 for the 108 clinical cases (see Table 8). Moreover, the width of the inter-quartile range of HoNOSI total scores was relatively narrow (from 6 or 7 to 11 points) for these three vignettes compared with a range of 12 points for the 108 clinical cases.

Table A7.2: *Distribution of HoNOSI Total Score for the 3 Vignettes*

Vignette #	n	mean	sd	min	max	Percentile				
						10th	25th	50th	75th	90th
1 - Aiden	81	25.7	5.4	11	40	20	23	26	29	32
2 - Emma	79	30.5	7.9	14	48	22	25	30	36	43
3 - Sophia	81	32.0	5.6	18	43	25	28	32	35	40

Cronbach's alpha coefficient for the HoNOSI total score was 0.66, 0.82 and 0.75 for each of the 3 vignettes respectively, indicating a high degree of internal consistency in how the scales are interrelated and is comparable to that reported for the 108 clinical cases.

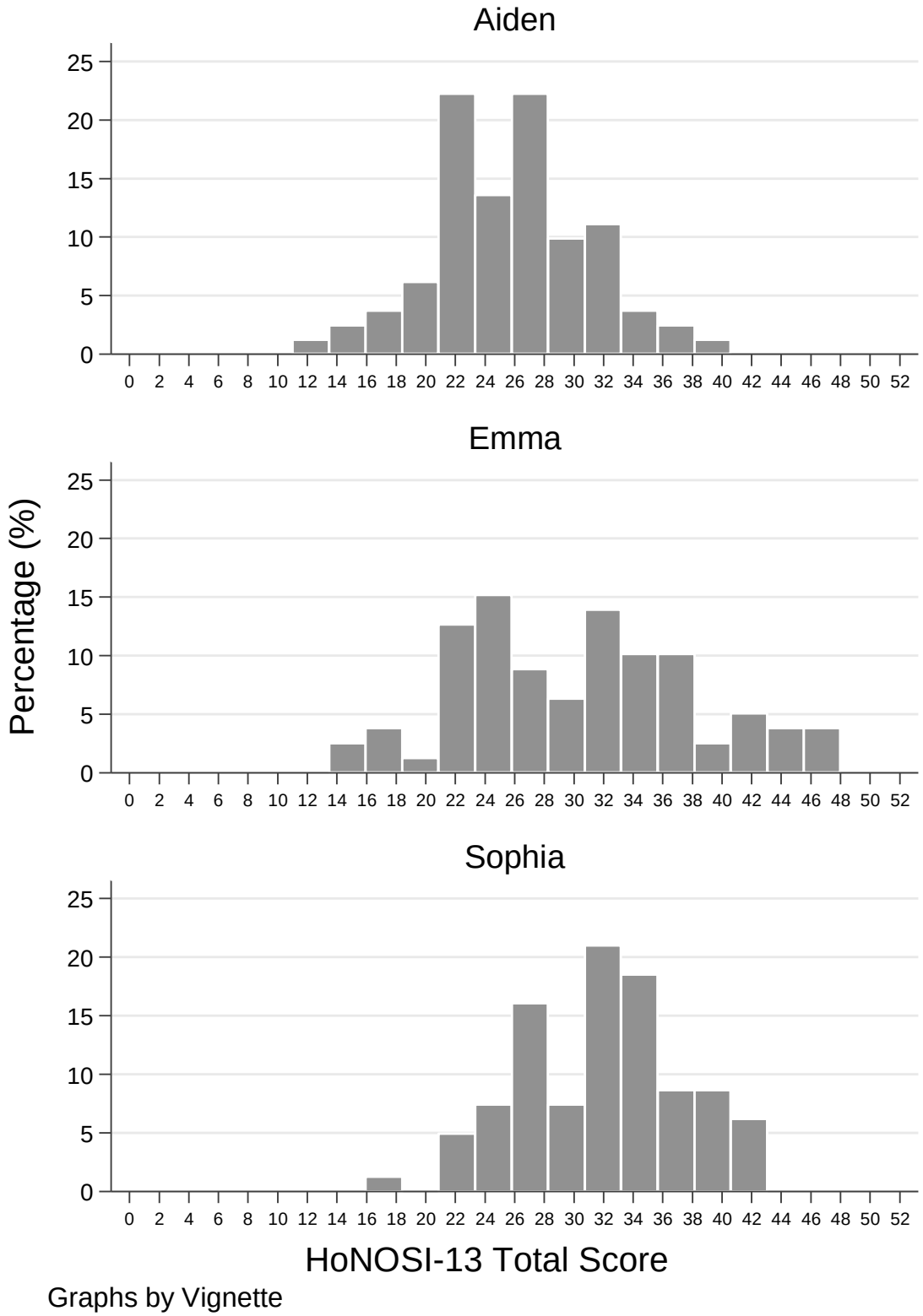

Figure A7.1: Distribution of HoNOSI Scores for the 3 Vignettes

Inter-rater Reliability of Case Vignette Data

In COSMIN, for measures that are constructed for ratings at an ordinal level of measurement (i.e., the rating categories indicate the relative ordering and not the actual "quantity", of levels of "clinical severity"), the "gold standard" for reporting is the weighted kappa (Mokkink, 2018).

Weighted kappa statistics incorporate adjustments designed to account for rating discrepancies. The effect is to provide relative "credit" for different ratings based on the number of category differences. For example, a discrepancy of one rating category, say a HoNOSI severity rating of 1 *Minor* vs. a 2 *Mild* is less important than a discrepancy of two rating categories, say a 1 *Minor* vs. a 3 *Moderate*, and so on.

The intraclass correlation coefficient (ICC) has also been used to evaluate inter-rater reliability. This model is appropriate within the COSMIN framework where the level of measurement is continuous (i.e., interval) rather than ordinal. For continuous scales, an interval difference of 1 point has the same meaning across the entire range of rating points (e.g., the 1-point rating difference between 1 *Minor* and 2 *Mild* is the same as the 1-point rating difference between 2 *Mild* vs. 3 *Moderate*, and the 1-point difference between 3 *Moderate* vs. 4 *Severe to Very Severe*.

ICC models also allow for estimation of separate effects for "absolute agreement" vs. "consistency of agreement". "Absolute agreement" concerns whether different raters assign the same rating to the same subject – that is, whether raters use the same number. For example, the extent to which different raters all agree that case x is rated 12 points and case y 14 points and case z is 17 points. "Consistency of agreement" concerns the degree to which raters are consistent in their ordering of different targets. For example, different raters agree that case x is less severe than case y which is less severe than case z.

In this study, inter-rater reliability was assessed using both weighted kappa statistics and intraclass correlations. For kappa, ordinal weights were applied. For intraclass correlations, a two-way random-effects model was estimated (i.e., this assumes that the sample of HoNOSI raters is representative of all HoNOSI raters). Intraclass correlation coefficients were estimated separately for absolute agreement and consistency of agreement. While the three metrics are valid estimates of inter-rater reliability, they differ in that weighted kappa is more statistically precise than absolute ICC, which is more precise than consistency ICC (i.e., more precise estimates have narrower confidence intervals). Moreover, more precise estimates are generally smaller than less precise estimates (i.e., weighted kappa statistics are smaller than ICC statistics).

COSMIN offers a standard for acceptable inter-rater reliability (i.e., IRR is at least 0.70). There are a number of other guidelines for describing the *relative* strength of agreement associated with inter-rater reliability estimates. The Landis and Koch (Landis and Koch, 1977, p. 165) guidelines are the most frequently used and are shown below.

< 0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost Perfect

Table A7.3 summarises the inter-rater reliability findings. Weighted kappa, absolute agreement intraclass correlation and consistency intraclass correlation coefficients are colour coded to show the different Landis and Koch descriptions (e.g., the light blue cells indicate coefficients in the range 0.00 to 0.20 – *Slight*, dark green cells indicate coefficients in the range 0.81 to 1.00 – *Almost Perfect*). Appendix 12 provides further detail.

As would be expected, weighted kappa estimates are generally lower than ICC estimates and ICC consistency estimates are generally somewhat larger than ICC absolute agreement estimates. It can be seen that across all three inter-rater reliability estimates, only two of the 15 HoNOSI Scales (Scale 4 *Problems with feeding and eating behaviour* and Scale 6 *Problems with physical illness or disability*) satisfy the COSMIN criteria of acceptable inter-rater reliability. Estimates were lowest for Scale 7 *Problems associated with regulation and integration of sensory processing* and Scale 2 *Problems with activity levels, joint and/or sustained attention*. Using the consistency ICC model, only two scales were *almost perfect*, one was *substantial*, five were *moderate*, six were *fair* and two were *slight*.

Table A7.3: *Summary of Inter-rater Reliability Model Estimates*

	HoNOSI Scale	Raters (n)	Weighted Kappa (95% CI)	ICC Absolute (95% CI)	ICC Consistency (95% CI)
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	81	0.21 (0.13-0.29)	0.32 (0.11-0.95)	0.38 (0.14-0.96)
2	Problems with activity levels, joint and/or sustained attention	81	0.07 (0.02-0.13)	0.13 (0.03-0.86)	0.19 (0.05-0.91)
3	Non accidental self-injury or lack of self-protective behaviours	78	0.45 (0.35-0.56)	0.59 (0.27-0.98)	0.64 (0.31-0.99)
4	Problems with feeding and eating behaviour	81	0.83 (0.78-0.89)	0.91 (0.72-1.00)	0.91 (0.73-1.00)
5	Problems with developmental delays	81	0.26 (0.13-0.38)	0.34 (0.12-0.95)	0.44 (0.17-0.97)
6	Problems with physical illness or disability	80	0.72 (0.63-0.82)	0.82 (0.55-0.99)	0.81 (0.53-0.99)
7	Problems associated with regulation and integration of sensory processing	78	0.03 (-0.01.-0.08)	0.06 (0.01-0.75)	0.09 (0.02-0.82)
8	Problems associated with sleep	79	0.32 (0.26-0.39)	0.44 (0.17-0.97)	0.47 (0.19-0.97)
9	Problems with emotional and related symptoms or over-controlled emotional regulation	81	0.10 (0.04-0.16)	0.17 (0.05-0.90)	0.23 (0.06-0.92)
10	Problems with social reciprocity	78	0.12 (0.07-0.18)	0.20 (0.05-0.91)	0.24 (0.07-0.93)
11	Problems with age appropriate self-care and environmental exploration	76	0.19 (0.12-0.26)	0.29 (0.09-0.94)	0.36 (0.12-0.96)
12	Problems with family life and relationships	81	0.28 (0.20-0.36)	0.41 (0.15-0.97)	0.45 (0.17-0.97)
13	Problems with attending care, education and socialisation settings	75	0.35 (0.23-0.46)	0.47 (0.18-0.97)	0.49 (0.20-0.98)
14	Problems with knowledge or understanding about the nature of the infant's difficulties	80	0.18 (0.09-0.26)	0.27 (0.08-0.94)	0.36 (0.12-0.96)
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	79	0.22 (0.16-0.28)	0.34 (0.11-0.95)	0.36 (0.12-0.96)
	Total HoNOSI	79	0.14 (0.09-0.19)	0.22 (0.06-0.92)	0.42 (0.15-0.97)

Inter-rater reliability estimates:

< 0.00	Poor
0.00-0.20	Slight

0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost Perfect

Discussion

This study was designed to test inter-rater reliability of the HoNOSI Scales. Findings showed that the level of inter-rater reliability was relatively poor. This was the case with all three metrics. Even though three different metrics were used to estimate inter-rater reliability, there were no overall differences as far as the COSMIN threshold.

Additional statistical analyses were undertaken to check that the design of the inter-rater reliability trial was sufficiently powered (i.e., 3 vignettes and 81 clinical raters). Findings from these analyses confirmed that the research design was sufficiently powered to test the COSMIN threshold of "adequate" inter-rater reliability. There are several other possible explanations for the poor level of inter-rater reliability.

First, there are likely effects concerning the nature of the vignettes themselves. As noted earlier, the three vignettes were relatively homogeneous with respect to the overall HoNOSI total score. As noted by Streiner et al., (Streiner, Norman and Cairney, 2014), the size of the inter-rater reliability estimates is related to the amount of variability among "targets" (i.e., variability in severity ratings across the three vignettes for any given HoNOSI Scale). As a concrete example of this concept, it is similar to people estimating different temperatures where three rooms are 26°C, 27°C and 28°C (ie., small difference; harder to estimate reliably) compared with three rooms of 15°C, 20°C and 25°C (ie., larger difference; easier to estimate reliably).

A simple way to test this idea in Study 2 is to plot the inter-rater reliability estimates as a function of the variability between raters. The coefficient of variation is the ratio of the standard deviation to the mean HoNOSI severity rating – smaller coefficients indicate lower variability between raters and larger coefficients greater variability. Figure A7.2 plots the relationship between the inter-rater reliability, as measured by the weighted kappa estimate, and severity rating variability, as measured by the coefficient of variation, for each of the 15 HoNOSI Scales. The figure shows, consistent with Streiner et al., 2014, a significant positive relationship between the size of the inter-rater reliability estimate and the amount of between rater variability ($r = 0.55$, $p < 0.05$).

Scales in the upper right region of the Figure, for example, Scale 4 and Scale 6 had both the highest kappa coefficients and greatest variability between severity ratings. Scales in the lower left region of the Figure, for example Scale 2 and Scale 9 had both the lowest kappa coefficients and least variability between severity ratings. Figure A7.2 shows that the majority of scales had low kappa and least variability.

The methodological flaw resulted in little variability between vignettes. This suggests that Study 2 may be limited in providing information on the inter-rater reliability of the HoNOSI.

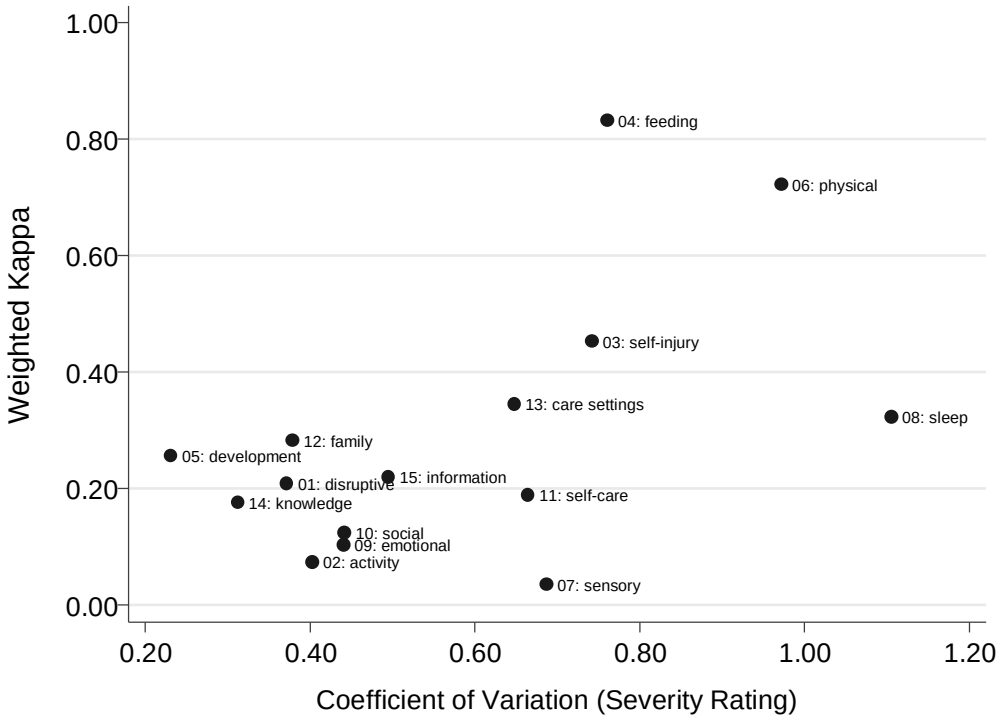

Figure A7.2: Relationship between strength of inter-rater reliability and HoNOSI Scale Severity Rating Variability

Second, there may have been issues related to the design and implementation of the inter-rater reliability field trial. For example, whether participants did in fact follow the protocol and rate vignettes independently, whether there were rating differences based on rating vignettes before or after rating actual clinical cases, etc. It may also be that rater clinical experience and professional background affected ratings – unfortunately, these clinician attributes were not available for the inter-rater reliability field trial.

Third, the validity of the HoNOSI Scales severity ratings may require further consideration. It may be the case that the HoNOSI glossaries, for say Scale 4 *Problems with feeding and eating behaviour* and Scale 6 *Problems with physical illness or disability*, provide more precise guidance for the clinical severity classes, than for example, Scale 2 *Problems with activity levels, joint and/or sustained attention*.

These ideas are not mutually exclusive and there may be multiple factors at play (e.g., issues concerning the "judges" (i.e., clinical raters) and issues concerning the "targets" (i.e., the 3 vignettes). Notwithstanding that and the speculative nature of some of these possible explanations, there is reasonable evidence to hypothesise that the poor levels of inter-rater reliability are likely an artefact of the vignettes themselves. In other words, methodological issues, particularly the small amount of variability between vignettes may have possibly masked the actual level of inter-rater reliability. A new inter-rater reliability field trial, for example, comprising a more heterogeneous set

of clinical vignettes, should be considered, especially given the positive and encouraging findings from the concurrent validity study.

Appendix 8: Case Vignettes - Inter-rater Reliability Pilot Study

Vignette 1: Aiden

Aiden is a 9-month old boy who is the first child of Kelly and John. Aiden was a long, complicated delivery requiring vacuum extraction. Aiden had an Apgar score of 4 on delivery.

Aiden is physically healthy. After an initial period of difficulty with feeding, he is now seen as a good eater. His parents report that his sleep is generally OK although it sometimes takes a while to resettle him if he wakes in the evening.

Aiden is impulsive and shifts his attention to most things that come into his field of view. He was described as becoming easily distressed over anything. Small changes in his routine can be very unsettling. Aiden still does not interact well with other children, screaming and hitting them. However, he is responsive to, and enjoys playing games with, both of his parents.

From the parents' report and from direct observation, when there are changes in his routine or when things do not go the way that he wants them to go, he becomes almost uncontrollable, screaming to the extent that it appears he is unable to breathe. When he is this distressed, he has been observed pulling his hair out. These behaviours are displayed regardless of who is caring for Aiden, either Kelly or John or their parents, with both sides of the family very supportive. John believes that outside of these frequent tantrums, Aiden is reasonably relaxed, responsive and happy.

Aiden has been slow to reach his developmental milestones. Aiden has some expressive and receptive language difficulties. Aiden is still not babbling and still does not appear to recognise familiar words. He still does not seem to recognise his own name.

Kelly and John have tried to put Aiden in day care so that Kelly can return to work, however Aiden became so distressed and interacted so poorly with other children that they decided not to proceed. Aiden's parents do not know what to make of his behaviour. They have also been confused by the conflicting advice they have received about what to do to help him.

Kelly is a 29 yr. old lawyer who has a history of anxiety, often triggered by stressful events. Previous treatments include SSRIs and CBT. She also has a strong family history of anxiety disorders. John notices that it seems to take Kelly a long time to fall asleep because she constantly worries about Aiden. If Aiden cries, Kelly leaps to hold him and sooth him. Kelly is worried that her attending so quickly is preventing John from participating in this aspect of Aiden's care. John was less worried about that, but did worry that over time this will increase Aiden's need for mum.

Vignette 2: Emma

The paediatric liaison team from a major hospital contacted the service with concerns about Emma, a 4 month old girl. Emma is currently an inpatient for some complicated medical problems. She has problems with her bowels and also requires nasogastric feeding. She has been diagnosed with a major developmental delay. A referral to disability services has been made. Emma is a twin and her brother is reported to be doing well. The family live in a small house but with enough room for all of them.

The pregnancy was difficult. Since birth Emma has been difficult to settle to sleep, though this seems to be OK over the last few weeks in hospital. Emma will often arch her back and push away when mum goes to hold her, and sometimes Emma acts like mum is not there at all. Emma appears to swing from periods of restlessness to being listless. When restless, she is interested in everything around her in the hospital. There is a concern that while her hearing is intact, she often seems under-responsive to noise. However, direct observation shows that she appears to be interested in everything around her in hospital.

Mum's expressed concerns were that she had not spent much time with her daughter while she was in hospital but she felt that was getting better. There is an older child and mum felt that the demand on looking after 3 infants under 2 years old was not being taken seriously. She thought Emma was difficult to handle with vomiting and needing lots of special individual attention.

Mum reports feeling that she does not feel close to her daughter and is frightened by what to do for her disability. Mum has post-natal depression. There is concern from the paediatric team that mum does not attend all of her appointments and there is a belief that child protection may need to be involved.

Mum is not sure what the father understands about what is going on but knows that they both think the hospital professionals are doing too much when they want other agencies involved. They understand that Emma has developmental delay but think that she may have some undiagnosed allergies that are creating the difficulties between her and mum. The hospital staff think that mum is having difficulties attaching due to Emma's problems and mum's depression. The father is quite involved with the children and the relationship between parents appears supportive.

Contact with mum revealed that she was less concerned about her daughter but was very concerned about any child protection involvement. She was very reluctant to allow the service to talk to any other services. Mum stated that her concerns were all medical and she felt pressured by the 'professionals' to talk with the CYMHS. Dad agreed with these concerns but also thought that there was a lot of pressure on his wife and that maybe some help might be useful.

Vignette 3: Sophia

Sophia is a 3 year old girl living with her mother. The mother of Sophia contacted the mental health service after a number of months at a childcare centre. Sophia constantly ran off at this centre and the staff raised their concerns about this with mum. At home, she also screams, throws objects around and often knocks things off the table. She takes little care when breaking things and only last week hurt herself when throwing cutlery off the bench. Mum was worried that she did not appear concerned. While mum sees Sophia as angry, she is also worried that she is sad a lot of the time.

Sophia was described as sleeping well. She has been diagnosed with asthma but this only occasionally interferes with her activities. Her eating is fussy and she won't sit still at the table. She often eats wandering around the room. Sofia will play for periods of time with dolls and her dolls' house. There are difficulties in keeping Sophia clean as she is very hard to get in the bath and needs constant reminding to wash her hands after going to the toilet.

Mum reported that Sophia had been physically and verbally abused by Sophia's father and it has only been in the last several months that Sophia had started to use more than single words. Child protection services had been involved when Sophia was about 2 for some weeks and they had informed mum that the father had a criminal record of sexual assault on young children. The parents are separated for over a year now. She was sure that the separation would help Sophia but is not sure what to do next to help her girl.

Mum feels very supported by her parents and her and Sophia live in a self-contained unit at the back of her parents'. At child care, Sophia was described as not aggressive to the other children but tends to ignore them and just do what she wants. She does not play with the other children and does not share any of the toys she is using with them. From the moment she arrives at childcare, she stands highly alert and the staff are sure she is waiting to run.

The child care staff had also discussed Sophia's speech delay and mum wondered if this was a result of Sophia's father. Mum is pretty sure that everything stems from Dad being abusive and is unwilling to consider other possible factors. From direct observation of her play, Sophia was noted to include dolls that could represent a father figure in the doll's house, although this figure was repeatedly thrown out of the house.

Appendix 9: Frequency Distribution of HoNOSI Scales by Vignette - Inter-rater Reliability Pilot Study

Table A9.1: *Frequency Distribution of HoNOSI Severity Ratings - Vignette 1 (n = 81)*

	Vignette 1 – Aiden	Severity Rating					
	HoNOSI Scale	0 (%)	1 (%)	2 (%)	3 (%)	4 (%)	9 (%)
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	0.0	1.2	6.2	71.6	21.0	0.0
2	Problems with activity levels, joint and/or sustained attention	9.9	19.8	24.7	42.0	3.7	0.0
3	Non accidental self-injury or lack of self-protective behaviours	8.6	0.0	38.3	43.2	9.9	0.0
4	Problems with feeding and eating behaviour	82.7	13.6	3.7	0.0	0.0	0.0
5	Problems with developmental delays	1.2	1.2	18.5	64.2	14.8	0.0
6	Problems with physical illness or disability	91.4	0.0	2.5	4.9	0.0	1.2
7	Problems associated with regulation and integration of sensory processing	33.3	9.9	17.3	30.9	6.2	2.5
8	Problems associated with sleep	14.8	55.6	23.5	6.2	0.0	0.0
9	Problems with emotional and related symptoms or over-controlled emotional regulation	22.2	8.6	9.9	55.6	3.7	0.0
10	Problems with social reciprocity	6.2	6.2	21	53.1	12.3	1.2
11	Problems with age appropriate self-care and environmental exploration	25.9	17.3	17.3	30.9	6.2	2.5
12	Problems with family life and relationships	8.6	18.5	45.7	24.7	2.5	0.0
13	Problems with attending care, education and socialisation settings	0.0	4.9	7.4	14.8	72.8	0.0
14	Problems with knowledge or understanding about the nature of the infant's difficulties	2.5	7.4	30.9	54.3	3.7	1.2
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	11.1	19.8	39.5	25.9	2.5	1.2

HoNOSI Severity Ratings:

0 - No problem

1 - Minor problem requiring no formal action

2 - Mild problem
3 - Problem of moderate severity
4 - Severe to very severe problem
9 - Not known or not applicable

Appendix 9: Frequency Distribution of HoNOSI Scales by Vignette: Inter-rater Reliability Pilot Study

Table A9.2: *Frequency Distribution of HoNOSI Severity Ratings - Vignette 2 (n = 81)*

	Vignette 2 – Emma	Severity Rating					
	HoNOSI Scale	0 (%)	1 (%)	2 (%)	3 (%)	4 (%)	9 (%)
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	21.0	4.9	29.6	34.6	9.9	0.0
2	Problems with activity levels, joint and/or sustained attention	2.5	1.2	29.6	46.9	19.8	0.0
3	Non accidental self-injury or lack of self-protective behaviours	72.8	8.6	8.6	4.9	1.2	3.7
4	Problems with feeding and eating behaviour	0.0	1.2	1.2	12.3	85.2	0.0
5	Problems with developmental delays	1.2	0.0	0.0	13.6	85.2	0.0
6	Problems with physical illness or disability	3.7	1.2	2.5	23.5	69.1	0.0
7	Problems associated with regulation and integration of sensory processing	7.4	11.1	35.8	34.6	11.1	0.0
8	Problems associated with sleep	27.2	27.2	28.4	13.6	2.5	1.2
9	Problems with emotional and related symptoms or over-controlled emotional regulation	8.6	6.2	11.1	60.5	13.6	0.0
10	Problems with social reciprocity	22.2	7.4	13.6	29.6	23.5	3.7
11	Problems with age appropriate self-care and environmental exploration	35.8	11.1	24.7	17.3	6.2	4.9
12	Problems with family life and relationships	0.0	0.0	12.3	60.5	27.2	0.0
13	Problems with attending care, education and socialisation settings	54.3	7.4	6.2	8.6	16	7.4
14	Problems with knowledge or understanding about the nature of the infant's difficulties	1.2	1.2	8.6	32.1	56.8	0.0
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	0.0	3.7	12.3	42	42	0.0

HoNOSI Severity Ratings:

0 - No problem
1 - Minor problem requiring no formal action
2 - Mild problem
3 - Problem of moderate severity
4 - Severe to very severe problem
9 - Not known or not applicable

Appendix 9: Frequency Distribution of HoNOSI Scales by Vignette: Inter-rater Reliability Pilot Study

Table A9.3: *Frequency Distribution of HoNOSI Severity Ratings - Vignette 3 (n = 81)*

Vignette 3 – Sophia		Severity Rating					
HoNOSI Scale		0 (%)	1 (%)	2 (%)	3 (%)	4 (%)	9 (%)
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	1.2	0.0	9.9	54.3	34.6	0.0
2	Problems with activity levels, joint and/or sustained attention	6.2	3.7	16.0	58.0	16.0	0.0
3	Non accidental self-injury or lack of self-protective behaviours	4.9	3.7	39.5	44.4	7.4	0.0
4	Problems with feeding and eating behaviour	2.5	7.4	55.6	33.3	1.2	0.0
5	Problems with developmental delays	1.2	1.2	7.4	63.0	27.2	0.0
6	Problems with physical illness or disability	17.3	44.4	37.0	1.2	0.0	0.0
7	Problems associated with regulation and integration of sensory processing	30.9	6.2	17.3	39.5	2.5	3.7
8	Problems associated with sleep	92.6	4.9	1.2	0.0	0.0	1.2
9	Problems with emotional and related symptoms or over-controlled emotional regulation	2.5	2.5	4.9	61.7	28.4	0.0
10	Problems with social reciprocity	1.2	0.0	6.2	46.9	45.7	0.0
11	Problems with age appropriate self-care and environmental exploration	2.5	1.2	24.7	46.9	24.7	0.0
12	Problems with family life and relationships	3.7	3.7	7.4	34.6	50.6	0.0
13	Problems with attending care, education and socialisation settings	8.6	4.9	25.9	35.8	24.7	0.0
14	Problems with knowledge or understanding about the nature of the infant's difficulties	1.2	2.5	38.3	43.2	14.8	0.0
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	11.1	22.2	25.9	25.9	13.6	1.2

HoNOSI Severity Ratings:

0 - No problem
1 - Minor problem requiring no formal action
2 - Mild problem
3 - Problem of moderate severity
4 - Severe to very severe problem
9 - Not known or not applicable

Appendix 10: Frequency Distribution of HoNOSI Scales by Vignette - Inter-rater Reliability Pilot Study

Figure A10.1: Distribution of HoNOSI Scales by Vignettes

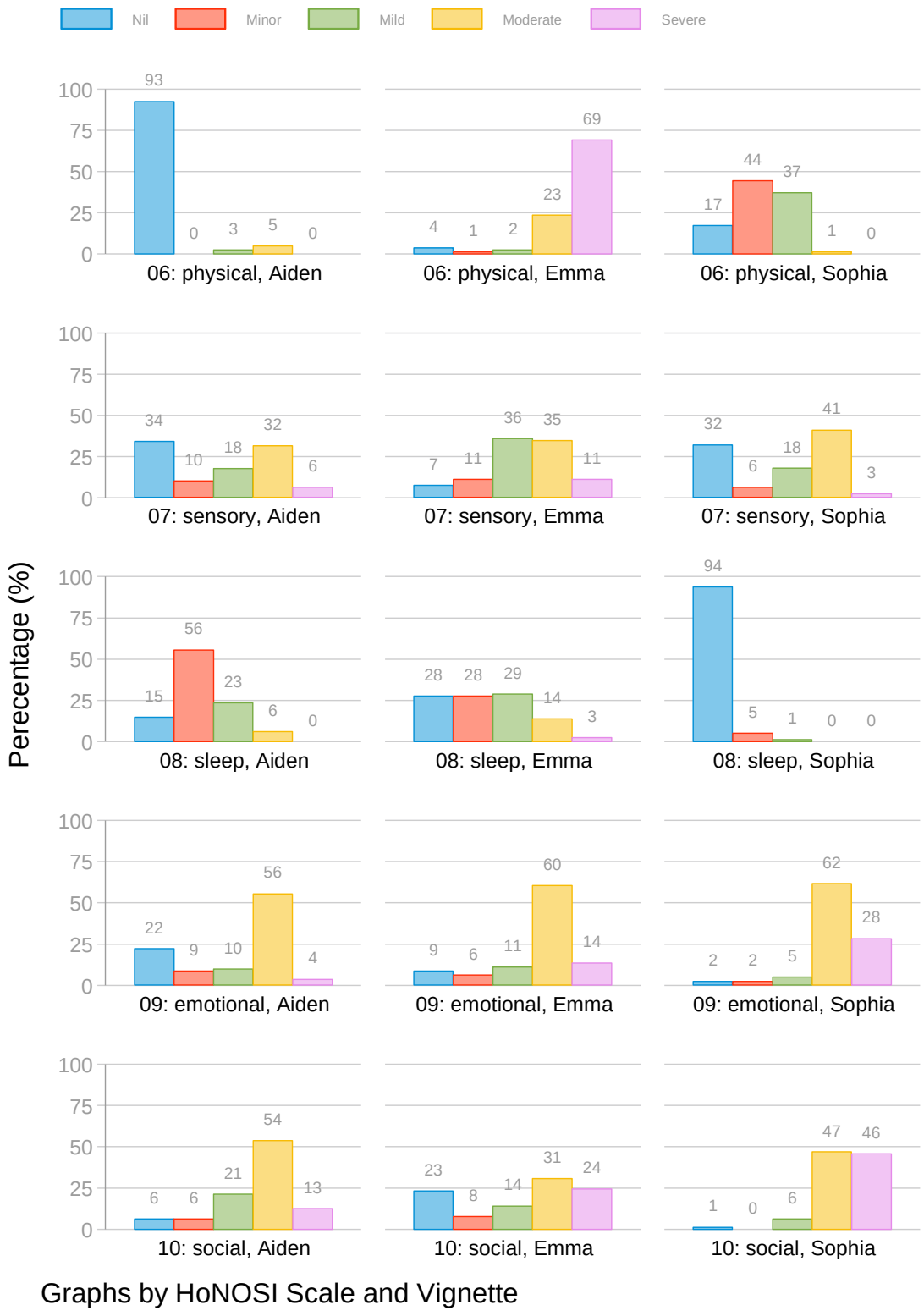

Figure A10.1: Distribution of HoNOSI Scales by Vignettes

Appendix 10: Distribution of HoNOSI Scales by Vignette

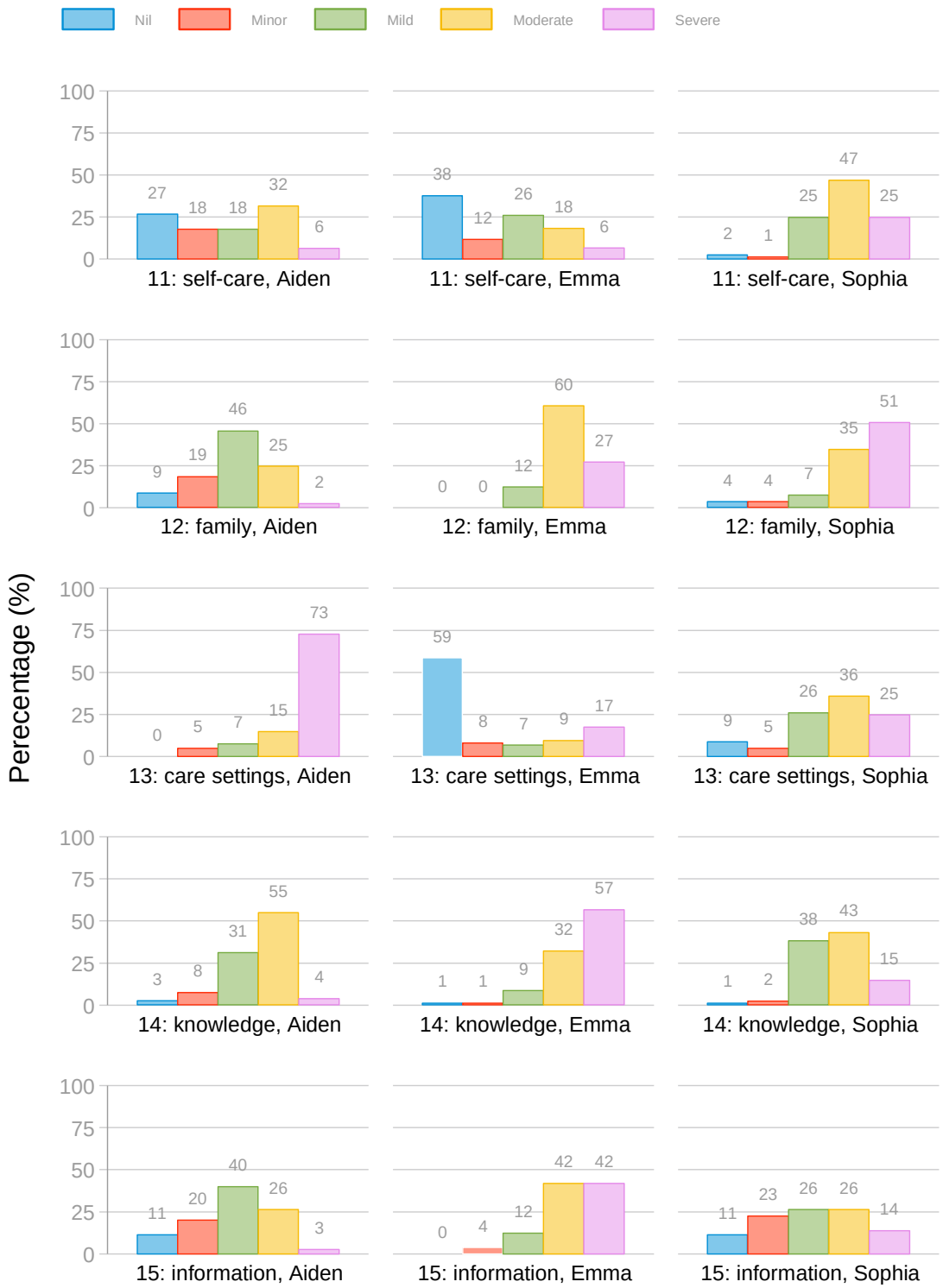

Figure A10.1: Distribution of HoNOSI Scales by Vignettes

Appendix 11: Missing Values Within Case Vignette Data - Inter-rater Reliability Pilot Study

Table A11.1: *Number of Missing HoNOSI ratings by Case Vignettes*

HoNOSI Scale		Vignette 1 (Aiden)	Vignette 2 (Emma)	Vignettes 3 (Sophia)	Raters (n)
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	0	0	0	81
2	Problems with activity levels, joint and/or sustained attention	0	0	0	81
3	Non accidental self-injury or lack of self-protective behaviours	0	3	0	78
4	Problems with feeding and eating behaviour	0	0	0	81
5	Problems with developmental delays	0	0	0	81
6	Problems with physical illness or disability	1	0	0	80
7	Problems associated with regulation and integration of sensory processing	2	0	3	78
8	Problems associated with sleep	0	1	1	79
9	Problems with emotional and related symptoms or over-controlled emotional regulation	0	0	0	81
10	Problems with social reciprocity	1	3	0	78
11	Problems with age appropriate self-care and environmental exploration	2	4	0	76
12	Problems with family life and relationships	0	0	0	81
13	Problems with attending care, education and socialisation settings	0	6	0	75
14	Problems with knowledge or understanding about the nature of the infant's difficulties	1	0	0	80
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	1	0	1	79
Total Score	HoNOSI	0	2	0	79

Appendix 12: Inter-rater Reliability Analysis of Case Vignette Data - Inter-rater Reliability Pilot Study

Table A12.1: *Weighted Kappa Estimates for Case Vignettes*

HoNOSI Scale		Raters (n)	kappa (wgt)	95% wkappa lower	95% wkappa upper	Strength of Agreement
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	81	0.21	0.13	0.29	Fair
2	Problems with activity levels, joint and/or sustained attention	81	0.07	0.02	0.13	Slight
3	Non accidental self-injury or lack of self-protective behaviours	78	0.45	0.35	0.56	Moderate
4	Problems with feeding and eating behaviour	81	0.83	0.78	0.89	Almost Perfect
5	Problems with developmental delays	81	0.26	0.13	0.38	Fair
6	Problems with physical illness or disability	80	0.72	0.63	0.82	Substantial
7	Problems associated with regulation and integration of sensory processing	78	0.03	-0.01	0.08	Slight
8	Problems associated with sleep	79	0.32	0.26	0.39	Fair
9	Problems with emotional and related symptoms or over-controlled emotional regulation	81	0.10	0.04	0.16	Slight
10	Problems with social reciprocity	78	0.12	0.07	0.18	Slight
11	Problems with age appropriate self-care and environmental exploration	76	0.19	0.12	0.26	Slight
12	Problems with family life and relationships	81	0.28	0.20	0.36	Fair
13	Problems with attending care, education and socialisation settings	75	0.35	0.23	0.46	Fair
14	Problems with knowledge or understanding about the nature of the infant's difficulties	80	0.18	0.09	0.26	Slight
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	79	0.22	0.16	0.28	Fair

Total Score	HoNOSI	79	0.14	0.09	0.19	Slight

Table A12.2: *Intraclass Correlation Estimates for Case Vignettes– Absolute Agreement*

HoNOSI Scale		Raters (n)	ICC	95% ICC lower	95% ICC upper	Strength of Agreement
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	81	0.32	0.11	0.95	Fair
2	Problems with activity levels, joint and/or sustained attention	81	0.13	0.03	0.86	Slight
3	Non accidental self-injury or lack of self-protective behaviours	78	0.59	0.27	0.98	Moderate
4	Problems with feeding and eating behaviour	81	0.91	0.72	1.00	Almost Perfect
5	Problems with developmental delays	81	0.34	0.12	0.95	Fair
6	Problems with physical illness or disability	80	0.82	0.55	0.99	Almost Perfect
7	Problems associated with regulation and integration of sensory processing	78	0.06	0.01	0.75	Slight
8	Problems associated with sleep	79	0.44	0.17	0.97	Moderate
9	Problems with emotional and related symptoms or over-controlled emotional regulation	81	0.17	0.05	0.90	Slight
10	Problems with social reciprocity	78	0.20	0.05	0.91	Slight
11	Problems with age appropriate self-care and environmental exploration	76	0.29	0.09	0.94	Fair
12	Problems with family life and relationships	81	0.41	0.15	0.97	Moderate
13	Problems with attending care, education and socialisation settings	75	0.47	0.18	0.97	Moderate
14	Problems with knowledge or understanding about the nature of the infant's difficulties	80	0.27	0.08	0.94	Fair
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	79	0.34	0.11	0.95	Fair
Total Score	HoNOSI	79	0.22	0.06	0.92	Fair

Table A12.3: *Intraclass Correlation Estimates for Case Vignettes - Consistency*

HoNOSI Scale		Raters (n)	ICC	95% ICC lower	95% ICC upper	Strength of Agreement
1	Problems with disruptive behaviour/irritability/under controlled emotional regulation	81	0.38	0.14	0.96	Fair
2	Problems with activity levels, joint and/or sustained attention	81	0.19	0.05	0.91	Slight
3	Non accidental self-injury or lack of self-protective behaviours	78	0.64	0.31	0.99	Substantial
4	Problems with feeding and eating behaviour	81	0.91	0.73	1.00	Almost Perfect
5	Problems with developmental delays	81	0.44	0.17	0.97	Moderate
6	Problems with physical illness or disability	80	0.81	0.53	0.99	Almost Perfect
7	Problems associated with regulation and integration of sensory processing	78	0.09	0.02	0.82	Slight
8	Problems associated with sleep	79	0.47	0.19	0.97	Moderate
9	Problems with emotional and related symptoms or over-controlled emotional regulation	81	0.23	0.06	0.92	Fair
10	Problems with social reciprocity	78	0.24	0.07	0.93	Fair
11	Problems with age appropriate self-care and environmental exploration	76	0.36	0.12	0.96	Fair
12	Problems with family life and relationships	81	0.45	0.17	0.97	Moderate
13	Problems with attending care, education and socialisation settings	75	0.49	0.20	0.98	Moderate
14	Problems with knowledge or understanding about the nature of the infant's difficulties	80	0.36	0.12	0.96	Fair
15	Problems with lack of information, understanding about services, or managing the infant's difficulties	79	0.36	0.12	0.96	Fair
Total Score	HoNOSI	79	0.42	0.15	0.97	Moderate

Appendix 13: Case Vignettes: Inter-rater Reliability Study

Vignette 1: Levi

This information relates to the last two weeks, unless otherwise stated.

Levi is a 9-month old boy who is the first child of Kelly and John.

Levi is impulsive and shifts his attention to most things that come into his field of view. He is often restless although he can be helped to settle with support. Levi does not interact well with other children; he screams and hits them. From the parents' report when things do not go the way that he wants them to go, he can become almost uncontrollable, with excessive screaming and lashing out at nearby objects. When he is this agitated, he has been observed pulling at his hair. Often, this results in multiple strands of hair being pulled out.

After an initial period of difficulty with feeding, he is now seen as a good eater. Levi was slow to reach some developmental milestones with some expressive and receptive language difficulties. Levi is still not babbling and still does not appear to recognise familiar words.

Levi is physically healthy. He appears to have no problems with integrating sensory stimuli. He is neither under nor overly sensitive to any sensory stimuli. His parents report that his sleep is generally OK although occasionally it takes a short while to resettle him if he wakes in the evening.

Outside of his anger tantrums, Levi is reported to be mostly relaxed and happy although occasionally is felt to be sad by his dad. His dad is not sure why Levi sometimes seems sad. Dad is a little worried and would like to keep an eye on this to ensure it does not worsen. He is not always socially responsive to his parents and others. For example, his parents are distressed that his response to games like peek a boo are patchy and that he often fails to engage in turn taking or imitation games. Compared to other children, he seems to have a moderate level of difficulty with social reciprocity. To his parents' surprise, he engages in his daily activities such as dressing, changing and bathing and is good at self-care.

Kelly and John appear to have a good relationship and readily keep Levi's needs in mind. They recently put Levi in day care so that Kelly could return to work. However, Levi's poor interactions with other children has led to him missing several sessions recently.

Levi's parents do not know what to make of his behaviour. They have also been confused by the conflicting advice they have received from family and friends about what to do to help him. Since coming to the service, they appear eager to change whatever is necessary, including their own responses, to help him. While still being unsure of why Levi is behaving the way he is, they both feel they have a better idea of how to manage his difficulties.

Vignette 2: Lily

This information relates to the last two weeks, unless otherwise stated.

The paediatric liaison team from a major hospital contacted the service with concerns about Lily, a 4 month old girl. Lily is currently an inpatient for some medical problems. She has structural problems with her bowels and surgery has been recommended.

The pregnancy was difficult. Lily has not shown any excessive irritability or oppositional behaviours. Lily occasionally can appear to swing from periods of mild restlessness to listlessness. There are no signs of any self-injurious behaviours. Feeding has always been difficult with Lily and she currently requires nasogastric feeding. The paediatric team and the parents are not concerned about any issues with her development or even milestones generally. Their concerns are about specific symptoms, noted below.

There is a concern that, while her hearing is intact and she does appear to attend to activity around her, she definitely has regular periods where she is under-responsive to noise and to voices. At these times, it appears as if she is overloaded with input, rather than being inattentive.

Since birth, Lily has had some minor difficulties settling to sleep, although this has resolved over the last month. Mostly, she appears appropriately interested in everything and everyone around her in the hospital and the team are not very concerned about this.

Lily occasionally seems to be anxious but this seems to be a minor problem and there is no concern about this. Lily does seem to gain comfort from her parents when they are present. She smiles and giggles and appears quite responsive to their gaze and engagement with her. Her eye contact is good with staff as well as parents.

Mum is concerned that she had not spent as much time with her daughter in hospital as she would have liked. There is a two year old sibling and mum felt that looking after him made it difficult to spend a lot of time with Lily. Mum has mild post-natal depression and there is concern from the paediatric team that mum does not attend all her own appointments. While the staff do not believe mum's depression is having a major impact on Lily, they do think there is some impact from her depression. It is clear that sometimes mum appears to have trouble keeping Lily's needs foremost in her mind.

Both mum and dad understand that Lily has serious health problems and that her under responsiveness to noise may be a consequence of this. Dad is quite involved with the children and the relationship between parents appears supportive. Mum stated that her concerns for Lily were all medical and she felt pressured by the doctors to talk with the CYMHS. Dad agreed with these concerns but also thought that there was a lot of pressure on his wife and that maybe some help might be useful.

Vignette 3: Chloe

This information relates to the last two weeks, unless otherwise stated.

Chloe is a 3 year old girl living with her mother. The mother of Chloe described that Chloe had spent a number of months at a local childcare centre. Chloe readily attends the centre. The staff at the centre had told mum that Chloe was developing very well in most areas. They told her that Chloe was a delightful girl although they also mentioned a couple of areas of concern. These concerns led Chloe's mum to contact the mental health service.

Both mum and the childcare centre noted that Chloe's eating is fussy and that this fussiness does not appear connected to any one particular type of food. Chloe doesn't sit still at the table and sometimes eats while wandering around. No-one considers this fussiness around food anything other than a minor issue. The staff noted that Chloe sits at the table when reminded although she may need reminding again in a few days. Neither the staff nor mum believed that Chloe was being defiant.

At childcare, Chloe was described as not aggressive to the other children and as playing well both beside, and with, them. She shares toys she is using fairly readily with them. Every now and then she becomes a little bit angry with other children when they take a doll that she is particularly using; although the childcare staff felt this behaviour was not a real difficulty. From the moment she arrives at childcare, she seeks out the other children. Chloe will play for periods of time with dolls and her dolls' house. She concentrates well when engaged in activities she enjoys.

She has been diagnosed with asthma but this only interferes with her activities occasionally. Chloe was described as generally sleeping well. However, she occasionally has some early morning waking but is able to be resettled by mum. Her sleep is usually back to her normal pattern within a day or two at most. There are no reports of sadness or anxiety.

There are difficulties in keeping Chloe clean as she is very hard to get in the bath and needs constant reminding to wash her hands after going to the toilet. This toilet hygiene issue was also reported by the childcare staff who felt that she needed more reminding in this area than they expected from most children her age.

Mum feels very supported by her parents. She and Chloe live in a self-contained unit at the back of her parents' house.

While separated from Chloe's dad, mum felt that they were not in conflict over Chloe. She felt that they cooperated well when it came to Chloe. This was pleasing to both of them. Chloe's mum reported that both she and Chloe's dad thought the eating habits had developed during the separation or shortly afterwards when both parents tended to let things slide, but the impact had been minor on Chloe over the past six months. Neither parent was greatly worried but both were happy to seek advice from the childcare centre and to follow their recommendation to talk to this service.

Vignette 4: Sandy

This information relates to the last two weeks, unless otherwise stated.

Sandy is a four year old boy referred to the service by the family's General Practitioner. Sandy sometimes plays well with his 6 year old sister and other children he knows. However, he is aggressive to some of the smaller children at kindergarten (preschool), pushing them over and throwing toys at them. The kindergarten staff would like help with his bullying as it is affecting other children. Only mild signs of overactivity were observed and reported. With appropriate support, he can slightly reduce his activity level. There were no reports of excessive risk taking or injuries.

Sandy was described as having concerning feeding patterns. He rejects particular foods and becomes emotionally upset if they are left near him for too long. The family accommodate this, although it does make an already tense household even more on edge at meal times.

Sandy's parents think his language development has been slow. He said “Mummy”, “Daddy” and “Nini” (sister) at about two years old. At three, he uttered short sentences such as “Sandy eat”, “Mommy come”. He has longer sentences but has problems with consonant sounds that his peers readily produce. His physical development has been appropriate with no apparent problems in self-care milestones although he has minor problems with washing and teeth cleaning. He appears physically well.

His sleep patterns are very disturbed and he does not settle easily. He wakes at least a couple of times each night, on most nights calling out for his parents. This often wakes his sister and the parents are worried for her. He often wakes with nightmares about monsters. Sandy's monster worries intrude into his routine. He is highly anxious most of the time and clings to his parents when meeting others. He is quite frightened and mum is unsure if this is just about the monsters. The monster fear worsens in unfamiliar houses. She often has to ask if they can check the other rooms before he settles somewhat. His sister finds this very embarrassing.

It is also very difficult to persuade Sandy to leave Dad and go to kindergarten. Sandy's separation difficulties have had a definite impact on his self-care. Sandy's anxiety, clinging and kindergarten absences all interfere with his interactions with other children. He can be disengaged and not enter any dyadic play. This sometimes settles after a couple hours and he then plays slightly more appropriately with other children. Mum worries about exacerbating his anxiety if she insists that he attends and dad is also reluctant to confront Sandy's worries by taking him to kindergarten. Sandy is absent for the majority of kindergarten sessions.

The family are under considerable financial and emotional stress. Dad lost his job months ago and appears moderately depressed. He spends most days either on the Internet, playing games, or drinking with a group of friends. Mum's family has a history of anxiety and she is concerned that Sandy does not develop the same problems. Both parents openly and angrily blame each other for Sandy's problems and this was both observed and reported to occur in front of both children. Both children are noticeably upset by this.

The parents are very worried. Father says that “Sandy is a little psychopath” and that he knows that people like Sandy often end up in prison or as drug addicts. The parents said they asked for a referral to this service because Sandy needs medication.

Vignette 5: Sally

This information relates to the last two weeks, unless otherwise stated.

Sally is a 3 year old girl living with her grandmother. Sally's grandmother contacted the mental health service with concerns about Sally's behaviour over the last couple of months at the childcare centre she attends. Sally constantly runs away from this centre and has hit staff who have tried to stop her pushing other children. At home, she often hits her grandmother; she screams, throws objects around and often knocks things off the table. Although Sally will play for periods of time with her dolls, she is very restless and there is no developing story in her play - the dolls have random roles. She often breaks things, not caring if she injures herself. Only last week she hurt herself badly again, when throwing cutlery off the bench. Sally's eating is a bit fussy, although her grandmother is not greatly worried about her nutritionally. She always eats wandering around the room.

It has only been in the last several months that Sally has started to use more than single words. Sally often becomes very distressed when she tries to express herself, an indication of her language delay. She has asthma and this moderately interferes with her physical activity. At childcare, Sally was described as being highly alert from the moment she arrives, seeming to be severely sensitive to noise and movement. Staff are very concerned about this reactivity, especially as her hearing and eyesight have been investigated and found to be functioning normally.

Sally was described as sleeping well. While grandmother thinks Sally is very angry, she also believes Sally is emotionally withdrawn and an extremely sad little girl. When playing with others, she does not take turns, she quickly becomes distracted and then ignores the other children. There are difficulties in keeping Sally clean as she is very hard to get in the bath and will continue to wear dirty underwear. Sally needs constant reminding to wash her hands after going to the toilet.

The grandmother reported that Sally had been physically and verbally abused by her father. Sally's mum can't believe that he has been abusive. The parents have been "partly" separated for over a year. Sally's mum is unsure of the type of relationship she wants with Sally's father. Child Protection Services had been involved with Sally since she was about 2 years old because of her father's criminal history of sexually assaulting children.

The grandmother and Sally's mum have discussed Sally's problems. They disagree on the cause, arguing over the role of the father's abuse and Sally's mum's belief that the interference of Child Protection is the key cause. The grandmother wants Sally's mum to have nothing to do with Sally's father, believing this would make everything alright. Her grandmother thought that having Sally live with her would help, but is not sure what to do next to help Sally.

Annex 1: HoNOSI Face Validity Study

HoNOSI Face Validity Study

CONTENTS

Executive Summary

The Health of the Nation Outcome Scale for Infants (HoNOSI) has been developed by the Australian Child and Adolescent Mental Health Information Development Expert Advisory Panel (CAMHIDEAP) as a routine outcome measure to support clinicians to improve the emotional and social well-being of children in the 0-47 month age group. A recent review of all relevant infant mental health outcome measures for infants in this age group identified a gap in outcome measures which address social, emotional and behavioural domains. The CAMHIEAP working group modified the existing Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA) and subsequently developed the HoNOSI in order to fulfil this need. The HoNOSI is currently in its draft stages of development. After an initial survey on the usefulness and relevance of the HoNOSI, the CAMHIDEAP Working Group identified the need for Face Validity Testing to inform further development. It is this Face Validity Testing phase of development which is the subject of this report.

In order to test Face Validity, a Focus Group study was designed with a representative sample of clinicians across Australia including psychiatrists, psychologists, nurses and other relevant allied health professionals who would be using the HoNOSI in routine practice.

In conclusion, focus groups showed that the HoNOSI fulfilled a much needed gap in Infant Mental Health Outcome Measurement for the 0-47 month age group as no suitable instrument previously existed. Overall, concluding thoughts to the overarching question pertaining to the purpose and usefulness of the HoNOSI and whether it fulfils its objectives, was unanimously positive in responses from all 6 out of the 6 face-to-face focus groups. They expressed a want and need to use it and welcomed the opportunity to have input and to make suggested improvements from a user perspective to inform next stages of development. Focus group findings and suggested areas of improvement are detailed in this report.

Background and Rationale

The Health of the Nation Outcome Scale for Infants (HoNOSI) has been developed by the Australian Child and Adolescent Mental Health Information Development Expert Advisory Panel (CAMHIDEAP) as a routine outcome measure to support clinicians to improve the emotional and social well-being of children in the 0-47 month age group. A recent review of all relevant infant mental health outcome measures for infants in this age group identified a gap in outcome measures which address social, emotional and behavioural domains. The National Outcomes and Casemix Collection (NOCC) Strategic Directions 2014 – 2024 review[1] of the routine outcome protocol in Australia identified a measurement gap for infants and pre-schoolers. The CAMHIDEAP working group modified the existing HoNOSCA[2] and subsequently developed the HoNOSI in order to fulfil this need. The decision to modify HoNOSCA rather than develop a completely new instrument at this point was partly informed by the desire to address this National need, the absence of alternative pragmatic options and the pragmatics of having an instrument that could occupy a place in the infrastructure constructed separately by all of the Australian state and territory governments for the HONOS and HoNOSCA family. The HoNOSI is currently in its draft stages of development. After an initial survey on the usefulness and relevance of the HoNOSI, the CAMHIDEAP Working Group identified the need for Face Validity Testing to inform further development. It is this Face Validity Testing phase of development which is the subject of this report.

Aim

The main aim was to test the Face Validity of the HoNOSI. This focus group study aims to elicit specialist mental health sector responses and responses from sectors in other relevant health services.

Method

In order to conduct Face Validity Testing, a Focus Group study was designed. It was anticipated that 4-6 participants would ideally be needed per focus group and that there would be one focus group per participating jurisdiction, for a representative sample of clinicians across Australia. Focus group members comprised a variety of HoNOSI end user representatives and consisted of psychiatrists, psychologists, nurses and other relevant allied health professionals who would be using the HoNOSI in routine practice.

Sample Size

It was anticipated that where feasible, focus groups would consist of 4-6 participants, with one focus group per jurisdiction, for a total of approximately (+/-) 30 focus group participants, or until responses reached the point of saturation. With a high level of interest and support from Working Group members and their colleagues in various jurisdictions, the anticipated sample size was met and exceeded, with a good representative sample across the country from most jurisdictions. By the end of the study, 5 of the 8 jurisdictions had registered their interest, coordinated staff members (end users) to be part of the focus group(s) in their jurisdiction, and made their staff available to participate in the focus group. Whilst 3 jurisdictions did not participate in the study, the number of participants was balanced out by one

1 National Mental Health Information Development Expert Advisory Panel (2013) Mental Health National Outcomes and Casemix Collection: NOCC Strategic Directions 2014 – 2024. Commonwealth of Australia, Canberra.

2 Gowers, S.G., Harrington, R.C., Whitton, A., Lelliott, P., Beever, A., Wing, J. and Jezzard, R. (1999). Brief scale for measuring the outcomes of emotional and behavioural disorders in children. Health of the Nation Outcome Scales for children and Adolescents (HoNOSCA). *British Journal of Psychiatry*, May: 174: 413-6.

jurisdiction volunteering two groups and some jurisdictions having larger numbers of participants. In total, 42 clinicians across 5 jurisdictions in 6 focus groups participated in the study.

Saturation

Whilst each group contributed some new information to the issues and themes in their focus group responses, it was deemed that the point of saturation was reached to a high level, given the law of diminishing returns in terms of new ideas or themes, with the number of focus groups completed. Whilst the sixth focus group did yield some new ideas for example, there was less new information gathered than seen previously in the first four groups. This can be evidenced in the focus group transcripts (Appendix 2), with the number of different perspectives diminishing as the study progressed, as expected. It was therefore deemed that 42 participants in six focus groups provided sufficiently diverse information in relation to the themes and issues for the HoNOSI to then inform further development of the measure.

Further to this, the non-focus group responses also yielded some new information though not additional overwhelming amounts, as expected. It must be noted that as per any study assessing the point of saturation, the ratio of new ideas to repeated concepts is expected to be greater at the beginning of the study as compared to the end of the study. This observation therefore is typical of that expected using this methodology; it is not a reflection of diminishing value of the ideas as the study progresses, rather a reflection of the order in which the focus groups were conducted.

Project Plan

The timeline for this project was January through June 2016. At the time of writing, this project is progressing well according to the schedule below.

Schedule

Planning meetings were held with AMHOCN. The study timetable is as follows:

<table>
<tr><td>January 2016</td><td>February</td><td>March</td><td>April</td><td>May</td><td>June</td></tr>
<tr><td colspan="4">Ongoing Project Co-ordination and Planning</td><td rowspan="2">Data Analysis</td><td rowspan="2">Write-up of results, Dissemination</td></tr>
<tr><td>Ethics, planning meetings</td><td>Ethics, Lit review, Administration</td><td>Recruitment and conducting focus groups</td><td>Conducting focus groups in each jurisdiction</td></tr>
</table>

Ethics

In terms of conducting the study, Ethics approval was not strictly required in terms of asking clinicians for their use on a draft instrument currently under development. As part of good practice however, Ethics approval was applied for through the NSW Institute of Psychiatry, who approved the study.

Focus Group Questions

The study utilised a mixed methods approach where possible, yielding both qualitative and quantitative data built on the original survey *Questionnaire on the HoNOSCI*[3]. For the qualitative section, the focus group participants were asked open-ended questions along the lines of those proposed below.

In order to obtain consistency in the pattern of responses across States and Territories (jurisdictions), significant project co-ordination and organisation of logistical issues was completed. This is illustrated in the project plan diagram (above) as an ongoing project activity within each participating jurisdiction, particularly during the focus group and data collection stages of the study.

A review of proof of concept regarding face validity was provided to the November 2015 HoNOSI working group, which is available upon request. The CAMHIDEAP meeting of December 2015 also reviewed and approved the proposal for the face validity testing. Draft questions proposed to be used for the focus groups, were as detailed below.

Analysis Plan

Which research questions will inform and how?

Focus group participants were asked to think about the following types of questions as they discussed each part of the HoNOSI, beginning with the Index of Scales (p3) and Key Principles (p4) and working through the scales sequentially from Scale 1 to Scale 15. Participants were asked to:

> Verbalise your thought process as you respond to the questions (Banna, Vera Becerra, Kaiser, et al., 2010).
>
> 1. Am I asking the right questions? (Frazer and Lawley, 2000)
> 2. Am I leaving anything out?
> 3. Are the questions brief?
> 4. Will the words be uniformly understood?
> 5. Are they too vague?
> 6. Are they too precise?
> 7. Have the criteria by which respondents must make a judgment been clearly stated?
> 8. How should the question be asked (Y/N, 5-point scale)?
> 9. Are items clearly worded? (Schneider and Rosenblum, 2014)
> 10. Do items adequately reflect the prescribed objectives of the instrument? (Schneider and Rosenblum, 2014)

Focus groups were recorded and data interpreted for the analysis. These findings are detailed in Appendix 2.

[3] It should be noted that initially the draft measure was known as the Health of the Nation Outcomes Scales for Children to Infants (HoNOSCI). However, this was later changed to the current title (i.e. HoNOSI) to eliminate any confusion with HoNOSCA.

Participants were also asked to keep in mind the overarching study question of "Is the instrument doing what it is intended to do?" OR another way to think about this "Is the instrument measuring what it is intended to measure?" They were informed that the group would then proceed to analyse and provide their insight into each Scale and at the end of the session, would be asked the overarching question. This was done for each group and the results are described at the beginning of the *Results* section.

Data used in the analysis

There were 6 face-to-face focus groups conducted in 5 participating jurisdictions: 2 in Victoria, and 1 in Sydney, Adelaide, Perth and Queensland. During the focus group in Adelaide, one of the emerging themes of *specialist knowledge*, which referred to the level of knowledge having an impact in the way the presenting observations were rated, led to the co-ordinator suggesting the facilitator conduct further in-depth exploratory focus group(s) or interviews to include Occupational Therapists, as per the example discussed at the time.

As some participants could not make it to the face to face focus groups, an additional 5 focus groups/interviews were conducted; 2 groups via teleconference and 1 person via phone interview from Adelaide and 2 individual Skype or in-person interviews from Sydney. These groups comprised Social Workers, Psychiatrists, Occupational Therapists and Psychiatrists.

Sample Size and Participants

In total, 50 people participated in the HoNOSI study (42 in face-to-face focus groups plus 8 non face-to-face responders). The face-to-face focus groups consisted of 42 participants across 5 jurisdictions with focus groups ranging in size from 5 to 10 (n=5,8,5,7,7,10). The aim was to have focus groups of 4-6 people, however the one large group of 10 people was necessary due to scheduling reasons. The teleconference groups (n=8 people) were smaller with between 1 and 3 participants, as expected, also due to scheduling and availability reasons. Additional written feedback was obtained from 2 participants in two jurisdictions as well as 1 person who had not participated in the focus groups but had provided written feedback.

Focus groups comprised of either one or more of the following clinicians: Psychologists, Occupational Therapists, Clinical Nurse Consultants, Clinical Nurse, Social Workers, Registrars, Psychiatrists, Speech Pathologists and a Medical Director. There did not seem to be a difference in responses between those working in Child and Adolescent Mental Health Services (CAMHS) versus Perinatal and Infant Mental Health Services (PIMHS) teams, rather, there was more of a notable distinction in the types of issues raised by those working with younger infants as opposed to older infants, in what was considered a wide age range (0-47 months) when considering the vast changes that occur in developmental stages during this age span.

How Codes were Created and Defined

The focus group recordings were transcribed and analysed to identify emerging themes within the data (Appendix 1). These data were then coded and with the creation of these codes, rules for applying the codes to data were defined (Appendix 1). As individuals were not identified within the study, themes were identified per group, as intended due to confidentiality reasons. Between-group comparisons were able to be made during the thematic analysis.

Teleconference focus groups and interviews were not able to be recorded. The information captured during these sessions was noted by the facilitator and written notes subsequently transcribed.

Outputs

There are two outputs expected – a draft report for submission to CAMHIDEAP and a draft peer-reviewed journal article, to follow.

Applied Thematic Analysis

In order to analyse the raw data resulting from the focus group responses, an Applied Thematic Analysis was conducted. This necessitated transcribing the original draw data into a transcription document and then coding the data and identifying relevant themes. The results of this can be seen in Appendix 1.

In order to make sense of this raw data, the transcribed data was then organised into thematic categories using the broader codes that had been applied. Codes are simply used to identify data pertaining to similar themes and as a tool to pull relevant themes together. The themes are the more important aspect of the analysis to consider and identify issues arising from focus group responses.

Results

Each focus group finished with an overarching question pertaining to the purpose and usefulness of the HoNOSI and whether it fulfilled its objectives. There were unanimously positive responses from all of the 6 face-to-face focus groups. Only 1 clinician of the 50 participating disagreed, noting that there was "no value in any of them", referring to measures in general, rather than the HoNOSI itself; "I fill in forms because my admin requires me to fill in forms". In terms of usefulness this clinician did state that the HoNOSCA was better than the HoNOS and that the HoNOSI in turn was better than the HoNOSCA. This comment was received from a participant in a teleconference discussion. All 42 face-to-face-focus group participants, plus the other 7 non-face-to-face participants (teleconference, etc) expressed either:

- Overwhelming positivity;
- positive in general, or;
- felt in general that it was good, but would be better with their suggested improvements.

These improvements and focus group participant responses are shown in the thematic analysis in the next section. Details of the focus group responses to the overarching question are shown here. Quotes are shaded, with square brackets denoting clarification or providing context for quotes. Ellipses indicate an alternate speaker contributing to the conversation.

Comments relating to the HoNOSI fulfilling a need were expressed as follows:

> I really like it, I feel it fills a vacuum, actually.

> Having an infant-specific outcome measure is great.

In terms of continuity:

> I think it's great that there's is a tool that's being developed that will fit alongside other tools that are accepted and regularly used and clinicians look at ratings, that makes sense to them. They're used to looking at five-point scales, used to knowing what that means.

> I would think this would be one of a suite of tools used by and discussed by a MDT

In relation to usefulness:

> Definitely facilitates discussion of infants' and young childrens' presentation amongst clinicians

General comments in relation to the overarching question (ellipses represent different speakers participating in the conversation):

I think it does
[agreement heard]
… Certainly does it better than what we had before, which was nothing!
… We've had to use HoNOSCAs that were nonsensical in parts, for infants

Partially to mostly
…*mostly* is a good word

With improvement, it would be much more useful than the HoNOSCA. Once you've finished with this one, can you review the HoNOSCA for Primary School ages children, 5-12 year olds?

Does some places, doesn't in other places… A lot of work has been done, and we really do appreciate it [from previous version]

Not without changes
…its getting closer
…of those changes are made, then yes
…if the over-two's it probably does it now, but for the little babies it needs modification
…the under 12 months
…We're really looking forward to using it
…Overall, we would be really pleased to have something like this that we could use
…it didn't actually take a massive amount of time; I think once you're familiar with it
...if we can make it work across the age range that we work with, it will be good.

I think it's got the potential to document the progress of infants and young children and make overt changes, but it comes back to the wide range of changes that occur between 0-47 months. I think that the tool has the capacity to pick that up, if the rater has the capacity to pick that up. My concern is that the tool requires some statement around who's going to be able to use it the way it was intended, because if you look at the group who put it together– highly specialised clinical people, but then I wonder about someone else picking that up

I would not say you would do this and it would inform treatment – it should inform assessment. For the other two aims I would say: tick, tick. [Person was looking at and referring to aims in the introductory section of the HoNOSI in responding to the overarching question.]

Initially it was clunky and cumbersome, but you'd be doing that day in and day out really.
… quite long, but once they get used to it…

The glossary is very comprehensive, but it doesn't necessarily address all of the nuances of this age range which is *such*…developmentally variable

I think is we looked at it as being a before and after measure and we're looking if there's improvements in their mental health, possibly not, given what we're talking about internalising and externalising stuff

I think it would capture that and I was thinking to put that to the test, so maybe when we have new patients [discussion with a staff member] you could do a before and after, because we haven't done it accurately enough to give that, for here.

Do you think it would work in a CAMHS setting?
…if it's not too targeted and narrow, that there is enough – the skill level is very different to make an assessment with young ones in CAMHS. At the moment it would be difficult for a CAMHS clinician
…but in a way, a scale like this could be helpful, because it is more explicit, so it can help to define out some of the abnormal behaviours

Two overall comments: it's a very wide age range: 0-9 or 12 months is different from 1-4 [years] or even 0-2. If you make the babies too wide, you're not going to catch the baby stuff, that's why I was suggesting 12 months

It's a good tool for general CAMHS to use to point them in the right direction of thought, what to capture, what to observe, what to think about, so I do think its useful
…would hone down the thinking – the focus needs to be in not just developmental difficulties in the child, but something parental

Applied Thematic Analysis Summary

Several key themes emerged from focus group responses. Importance was determined by the number of groups that talked about the issues, how much discussion they had around them and how strongly they felt about and expressed their views on the issue. Themes were not counted nominally nor ranked in terms of the raw value number of mentions, as a less important theme may have been mentioned many times, but that does not necessarily mean it was more significant. For example, clarity over wording and expression arose continuously, but the message over *Reference to Developmental Stages* and that of *Overlaps and Crossovers in Scales* and where/how to rate observations, was of greater concern than *Clarity* in the descriptions and ratings text, which was seen as an easier issue to resolve. Therefore, key themes identified in order of importance (roughly estimated from qualitative data as per Appendix 2) from focus group responses included:

1. Reference to developmental stages
2. Overlaps and crossovers in scales
3. Specialist knowledge
4. Global themes with subthemes of:
    a. Structure

   b. Length
   c. Consistency
   d. Clarity – global to the instrument
   e. Sources of Information
   f. Global Score
5. How to rate
6. Clarity - specific to descriptions or ratings

The most significant theme that emerged strongly and repeatedly was a need to have a reference guide to developmental stages in relation to the age of the child. This was closely followed by overlaps and crossovers in scales. Specialist knowledge also rated highly as a recurring theme. Global themes in relation to the HoNOSI in general related to the structure of the instrument, concern over its length, consistency across scales and many points of clarity that could be applied across the instrument; as opposed to applying to individual scales. Sources of information to use when rating and the notion of a global score also emerged as themes.

Questions and issues over how to rate observations were a frequently occurring theme, followed by much detail on clarifying either the descriptions and/or the ratings, for each individual scale. These themes are described in further detail in the following section. Table 1 shows a summary of the themes as they relate to each Scale.

Table 1: Overview of the number of Scales affected by each theme

Scale	Developmental Stages	Overlaps and Crossovers	Specialist Knowledge	How to rate	Clarity
1	✓	✓	✓	✓	✓
2	✓			✓	✓
3	✓	✓		✓	✓
4				✓	✓
5	✓	✓	✓	✓	✓
6	✓	✓			✓
7		✓		✓	✓
8	✓	✓	✓	✓	✓
9		✓			✓
10		✓	✓		✓
11	✓	✓		✓	✓
12		✓		✓	✓
13		✓		✓	✓
14	✓	✓	✓		✓
15		✓	✓	✓	✓

Theme 1: Reference to Developmental Stages

The theme that weighs most heavily as evidenced by focus group responses, which was independently raised, discussed and recurred frequently across 5 of the 6 groups, was that of the need to provide a reference to developmental stages in order to be able to utilise the instrument and rate observations accurately and consistently. It was suggested that it would be most useful to have a developmental reference guide included in the description of each scale which makes reference to the various developmental stages according to the age of the infant.

This stems from the rationale that the developmental stages vary so greatly across the 0-47 month age range, that less experienced users need a guide as to what is "normal´ in a 3-month old, compared with a 12 month old, compared with a three-year-old, for instance. Examples of what is *typical* for each developmental stage was repeatedly requested. The greatest level of concern in knowing how to rate developmental stages are at less than one year (at 3, 6, 9 and 12 months) versus 2 years versus 3 1/2 years of age.

Detail focus group responses are found in Appendix 3.1. Key quotes that illustrate these concerns are as follows:

> one of the issues will be what a 3 month-er can do, is going to be very different to what a 47 month-er can do, so therefore the person doing it needs to understand the difference between the two and I'm not clear that will happen

> …what does it look like in a 3 ½ year-old versus a 2 year old

> …I think you're right as the rate of change is so enormous – 3 months versus 9 months or one year, so you've got to know

Requirements are detailed with suggestions:

> a couple of sentences under each one [each scale] for infants under 12 months, or under six months

> Examples specifically broken down into different ages

> I think what we need is for people who *don't* do a lot of work with infants and would be expected to – there needs to be some kind of reference of what's to be expected for normal development – that's just my experience. "What is an exaggerated positive or negative emotional response?" How do you define exaggerated? What are the parameters around that for someone 0-3 months, 3-6, 6-9 – there's a lot that goes on in terms of very early development

This theme also related to theme 3, that of *specialised knowledge;* participants felt that if they were not experts in the field or were not familiar with using the instrument, they were not confident to rate their observations accurately.

Theme 2: Overlaps and Crossovers in Scales

The second of the three most significant themes was in relation to the length of the instrument, overlaps between scales, concerns of where to rate constructs, crossovers between scales and suggestions to delete certain scales.

Details of overlaps and crossovers are found in Appendix 3.2. The first column indicates the Scale in question or where there are crossovers and overlaps of two or more scales. During focus group discussion, 5 out of 6 groups raised issues about overlaps and crossovers and a different 5 out of 6 groups also raised issues about the length of the HoNOSI.

In working through the scales and discussing the overlaps and crossovers, one group suggested deleting scales 3,7,11,12, and 13, as their constructs could be accounted for within other areas of the HoNOSI. There was general acknowledgement that due to the HoNOSI draft having been derived from the HoNOSCA, it was understood why the current draft had 15 scales mirroring the HoNOSCA. However, it was discussed at length that this was not necessarily the best way to rate the 0-47 month age group and that the general structure could be substantially improved in terms of relevance, whilst also partly resolving the issue of length and time taken to complete the HoNOSI. Illustrative quote are as follows; for full details, refer to Appendix 3.2.

Regarding Scale 2 structure:

> Why put activity and attention together? You may have problems in one or the other domain and clinically if you have both, it's much more serious. I wonder if we should have two scales for those? I'm thinking about trauma and developmental disorders. I am thinking of a kid who's got a very high activity level and very poor attention as opposed to someone who's got an active temperament but can sustain attention.

Where to rate:

> In general, I think there is a lot of cross-pollination over the next three items [Scales 11,12,13]; how environmental exploration interplays with family life, interplays with opportunities for socialisation, and to some extent I think it will be sort of artificial how we pick those things

There were recurring comments about Scales 1 and 9:

> Dysregulation should probably come under Scale nine… So just extract the dysregulation symptoms... and put them into the examples for the emotional. The dysregulation is described in the descriptor in a non-aggressive way, and then also in an aggressive way

> Is it a behaviour scale?
>
> … I still think it should be separate [a separate scale – Disruptive behaviour separate scale to dysregulation]

In Relation to Scale 2:

> Again we were wondering about two scales here; the problem with activity levels, separate from joint and/or sustained attention

... Yeah, yeah, that really stands out doesn't it? There so much to take in.

Crossovers with Scales 1 and 7:

I think this is hard to tease out from Scale 1 for infants [probably until they are talking, or at least the under-one's], because you're seeing the same thing

…could fit in hyperarousal – where does that fit?

…exaggerated startled responses, or floppiness – where does that go with emotional regulation, as well

…its close, isn't it?

Discussion on Scale 13 and whether to delete or split scales:

Socialisation is elsewhere and Anxiety is elsewhere so you actually don't need it

…so ditch it

…and you could split one of the other two

… Not have that one but an elaborate on one of the other ones – developmentally does not fit with age group [school and backpacks mentioned here]

The above are some examples; for full details on this issue, refer to Appendix 3.2 and Appendix 1.

Theme 3: Specialist Knowledge

This theme describes the participant responses relating to where participants felt that they needed more "education or information" to feel like they were completing their ratings accurately. This theme is related to the first key theme of *Reference to Developmental Stages* as participants often describe feeling as though they needed specialised knowledge and expertise, particularly for the younger age groups in rating the scales. The overarching issue is that participants felt that a less experienced clinician may not be able to complete the HoNOSI accurately and they felt they needed more guidance around what to look for and that examples of what to expect would be helpful.

Specialist knowledge also referred to obtaining the opinions of specialist clinicians such as a paediatrician, speech pathologist, or occupational therapist, depending on the scale and constructs to be rated. Comments are detailed in Appendix 3.3. Examples include:

> I suppose the question is: How much knowledge is assumed about things outside the expertise of the rater, like, so for example, there's one of the scales that surrounds…speech and things

> My concern is that the tool requires some statement around who's going to be able to use it the way it was intended, because if you look at the group who put it together– highly specialised clinical people, but then I wonder about someone else picking that up

> I looked at this and thought I need a speech therapist in a way, to have a really good look at this area… People may think they've got a language problem when they've got an articulation problem [motor]

> I doubt this will be used for children under two, in the present situation… The point is you've got to make it understandable by someone who doesn't very often use it

> What is normal eye gaze – infants do turn away to regulate, come back after regulating include the word *sustained* avoidant gaze

> The description is not going to point people in the direction enough, if they're not a specialised service… The reality is, across Australia, most people would be generalists [referring to infants]

in relation to gaining familiarity, respondents also had this to say:

> I think there's learning effect – the more you do these, the more it sorts itself out; where to rate this, or where to rate that.

Theme 4: Global Themes

The three themes that are described above outline the most significant themes in terms of the importance placed on them, amount of discussion and frequency of response related to them, during focus group discussion. Several other global themes were identified and these have been categorised here in terms of *subthemes*. Global subthemes relate to issues that apply to the instrument globally and include issues related to:

4.1 Structure
4.2 Length
4.3 Consistency
4.4 Clarity – global to the instrument
4.5 Sources of Information
4.6 Global Score

4.1 Subtheme: Structure

Structure emerged as a theme as it was frequently discussed that the HoNOSI closely mirrors the HoNOSCA, where this may not always be appropriate in terms of structure. It was described as a "forced scale" in one instance.

Additionally, multiple groups have described the language and flow of the scale as seeming as though it was written by different people and needs to be made more consistent in terms of expression and style. Details can be found in appendix 3.4.1. Examples include:

> It feels a bit like a forced scale in some ways, its sort of adapted obviously from the HoNOSCA. You do see occasions of self injury, probably more in toddlers, probably less scalable I think than this

> That for me feels like its been put in because its in the HoNOS and the HONOCSA and it really isn't applicable

4.2 Subtheme: Length

Each focus group commented on the length of the questionnaire. The negative responses relating to length included discussion around the amount of time it would take to complete, concerns over the practicality of using the instrument due to its length and the fact that it is derived from the HoNOSCA and therefore not all of the Scales may be necessary.

Conversely, the positive responses in terms of length were that descriptions where necessary and useful for less experienced users, who found the level of detail and description helpful in guiding them on how to use the HoNOSI. The other take-home message was that once clinicians were familiar with it and were using the instrument frequently, it became a lot quicker to complete. Details are given in Appendix 3.4.2. Examples include:

> Initially it was clunky and cumbersome, but you'd be doing that day in and day out really.
> … quite long, but once they get used to it…

> The glossary is very comprehensive, but it doesn't necessarily address all of the nuances of this age range which is *such*…developmentally variable

> On the verge of what would be considered long in the include box
> … The shorter, more succinct, the more practical it will be in day-to-day use

4.3 Subtheme: Consistency

The theme of Consistency was of less concern compared to some of the other themes, but did emerge as a theme from multiple focus groups. Comments surrounded the language and style changing and needing to be made more consistent. Details are given in appendix 3.4.3. Examples include:

> Wondered if each scale is written by someone different; need to make it more consistent

> It looked like there might have been different people describing it and the language changed, so if there could be one editor using the same language all the way through

> It seems like certain specialists have contributed to the development of each of the scales… You get the sense there are slightly different styles, a little bit more detail in one and not the other… Ideally would have the same flow of language and similar level of detail for each scale

4.4 Subtheme: Clarity – Global to the Instrument

Clarity was also a relatively less significant issue which could be resolved with extra explanation or description. Examples include why certain scales were included (eg 14 and 15), clarifying that the observation and not the cause needs to be rated (or making this more explicit), and more specific examples around clarity and consistency in the ratings. Full details can be found in Appendix 3.4.4.

4.5 Subtheme: Sources of Information

This subtheme related to questions around how much weight to put on the clinician's observation versus parent report, where to rate certain cases and issues around collecting information from various sources. For a full description, refer to Appendix 3.4.5. One comment related specifically to ratings was that the participant liked the ratings to be kept simple while the bulk of the description is in the glossary above it.

> I prefer this type of rating where you obviously have to refer back to the explanation, rather than the explanation being embedded in the rating; I think that's easier

4.6 Subtheme: Global Score

This theme related to an overall assessment and the concept of a global or summary score at the end of the document. It was not raised frequently and is included here as it was a unique theme. See Appendix 3.4.6 for details.

Theme 5: How to Rate

This theme centred around which scale to use, the order of scales and whether they needed to be completed in sequential order and many questions surrounding where to rate which constructs. Comments referred to certain contexts, what is included or excluded, the severity of the ratings and what

they mean. Comments on how to rate raised the need for more clarity around the descriptions as well as the ratings themselves. See Appendix 3.5 for details. For example:

> I think I went to environmental contexts…I just said this could be very difficult to rate outside the caregiver context and then may require the assessor to try to interact with the infant, which, I do know that you guys, but I tend to think of not directly interacting with the infant but to observe the interaction…not sure whether this is now asking me to interact with the infant to see if I would get a different response from the infant to what I observe with their caregiver, so what I've said is this could be very difficult to rate outside of the caregiver context and then may require the assessor to try to interact with the infant to assess, that is assess with the caregiver and then assessment the clinician providing activity and appropriate support to see if the infant does things differently

Theme 6: Clarity - Specific to Descriptions or Ratings

Many of the above issues could be resolved, at least partially, with more clarity in the descriptions and ratings. Every group had much to say and many suggestions to make around the wording, interpretation and meaning of the descriptions and ratings of each scale. Appendix 3.6 contains a copy of each scale for easy reference, followed by focus group responses to that specific scale, for consideration. Many of these are "easy fixes" with suggestions provided by the clinicians on how to make the wording, the rating, the title or the Scale clearer. See Appendix 3.6 for details. For example:

> [Scale 1:] I had a particular issue with the word *settling* in here because it deliberately excludes sleep and I think that lots of people, with infants, when they talk about settling, what they talk about is settling them to sleep, which is actually addressed in a different scale, so you may want to look at another way of expressing because what I think they mean is being able to calm down

> [for Scale 10 Title, add:] *and interpersonal sensitivity*

> [Scale 12, Rating 1] *Parental reflective capacity* needs to be well defined. Do untrained people know what this is? Spell out more.

> [Scale 2, Rating 3:]– *appropriate supports* - I am not sure what is meant by appropriate supports; is this referring to parental supports, or what sort of supports are appropriate?

> [Scale 5, Rating 2] *mildly severe*

References

Banna, J. C., Vera Becerra, L. E., Kaiser, L. L. and Townsend, M. S. (2010). Using qualitative methods to improve questionnaires for Spanish speakers: assessing face validity of a food behavior checklist. *Journal of the American Dietetic Association,* 110(1)*,* 80-90.

Frazer, L. and Lawley, M. (2000). *Questionnaire design and administration. A practical guide.* Brisbane: John Wiley and Sons, Australia, Ltd.

Schneider, E. and Rosenblum, S. (2014). Development, reliability, and validity of the My Child's Play (MCP) questionnaire. *American Journal of Occupational Therapy,* 68(3)*,* 277-285.

Appendix 1: HoNOSI Face Validity Applied Thematic Analysis Code Book

Code	Code Description	Theme	Theme Description
Index	Pertains to Index section	Straight forward	Self Explanatory
Education and Training or instruction	Pertains to discussion on themes regarding - more knowledge of how to rate or how to use the scales -what the scales and rating mean in terms of applying this to an infant or child -how to recognise relevant observations against the norm for that age or developmental stage -guidance or further instruction required -subdivide age range	Specialist knowledge	-Requires expert or well-practiced knowledge in order to properly complete the scales -not seen as suitable for general clinician use -needing more education or information -pertains to comments on feeling of confidence and competence of the rater (clinician) -familiarity with using the scale -how much knowledge is assumed
		Sources of information	Rated by clinician and/or others
		How to complete	-Order of completion -Which scale to use -Where to rate what -How to rate
		Purpose	-how the tool informs assessment, not intervention
Rating	Pertains to rating (scores 0-4)	Global score	Overall score for the instrument
		Clarity	-Interpretation of meaning -How to rate
		Scores	-Size of gap between scores
		Additional tools	Refers to using additional tools for the purpose of assessing accurately
Global or Scales	Pertains to global/overall comments on the	Length	Of the instrument

Code	Code Description	Theme	Theme Description
	instrument; comments may pertain to all scales, or two or more scales		
		Usefulness	Of the instrument
		General	Comments in general about the scales or instrument
		Timeframe	The rating period
		Wording	Where clarity is required
		Structure	Structure of scales -amalgamating scales or items on scales -deleting scales -integrate scales -merging scales -splitting scales -consolidate scales
		Consistency	Across scales
		Continuity	With other tools such as the HoNOSCA
Wording	Change wording or language to make the meaning clearer	Clarity	Clarity needed – of meaning of words -scale -purpose of scale -what is being rated
Description	Termed the “description” in focus groups. Pertains to what the working group calls the “glossary” above the ratings table for each scale.	Clarity	-needs clarity -make meaning clear -interpretation -suggestions of how to improve the description
Title	The heading of each scale	Clarity	Title or purpose of scale needs clarity

Appendix 2: Transcription of Issues from Focus Group Data

For each focus group session, the facilitator had prepared for the purposes of consistency, a standard introduction which included the following talking points:

- Welcome/ thank you/introduction, methods and time length expected
- HoNOSI – developed as a routine outcome measure to support clinicians to improve the emotional and social well-being of children and adolescents 0-47 months
- For use in PIMs/CAMHS
- History- re HONOSCA – brief history
    - developed in the UK in late 1990's
    - trialed by Australian CAMS 1999
    - 2003 – part of the NOCC, implemented across Australia
- Where we are now
    - CAMHIDEAP – need for 0-47 months
    - Identified a gap in outcome measures
    - Working group has continued to modify the measure
- Participation audience
- Information Statement
- Consent Form
- Recording – verbal consent
- Questions – types of questions to think about as we discuss each Scale
- Written feedback also welcome
- Information gathering session

As this was done for each group, the transcriptions below start post-introduction or at the time of starting the recording (after verbal consent for recordings was obtained).

The overarching question/concept to think about is:

"Do items adequately reflect the prescribed objectives of the instrument?" or put another way: "Does this instrument do what we need it to do?"
"Does it measure what its intended to measure?"

Group 1

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
p3 – Index of Scales	Index	Straight forward	"Pretty straight forward"
P4 – Key Principles	Education and Training or instruction	Specialist knowledge	"I circled the one about *Ratings are informed by familiarity with, and a good understanding of, infant and child development* mainly because I think, well I suppose *familiarity* and *a good understanding of,* is not like, expert knowledge,...but that's the area I personally would like more, personally, more education or information about to feel…that I was doing this accurately…so like nearly a qualifier in that there'll be some areas where you feel like I'm competent to comment on, say *social reciprocity* where I'm not competent to comment on perhaps *the impact of or* maybe an early childhood nurse

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			who specializes in sleep would be competent to comment on Scale 8's *Problems associated with sleep*, so like having some, you can't be expected to go across them all but knowing that aohww [unsure sound], I think that's what you're looking at, where you don't have the knowledge in some areas, so you're a little bit making a judgement call instead of having an informed clinical decision on some scales, unless you're a developmental paediatrician in this area…so as any professional who would be doing it, whatever your skills would be, so as the Clinical Psychologist who worked with naught to four, *Scale 10, Problems with Social Reciprocity*, I'd go 'Yep, I'm competent to be a bit more able to assess on that' whereas [name omitted for confidentiality], you gave an example this morning of after two years of age, I don't know what's normal and what's not and there's no way to qualify, I think that's an issue.
P4 – Key Principles	Education and Training or instruction	Specialist Knowledge	I suppose the question is: How much knowledge is assumed about things outside the expertise of the rater, like, so for example, there's one of the scales that surrounds feeding, not feeding, speech and things, where it says that it should be rated even if there's not been further assessment by a speech pathologist or paediatrician, so it is asking for your, if you have a level of concern I suppose, so I guess its just a matter of maybe making clear in those principles that its, yeah, what level of expertise are anticipated with people rating because its kind of quite clear from the initial introduction that its targeting people who work with infant mental health primarily, but I guess, yeah, that that's going to be a range of people.
P4 – Key Principles	Rating	Global score	I guess what I'm not familiar with is, um, the document that I've downloaded, it doesn't actually show me what the outcome of the rating is, so (another person:) – You meant like a score out of… (original speaker:) – Yeah, at the end, yeah in terms of the, so the glossary doesn't like, um, and look I'm not really familiar, it's a long time since I've used the HoNOS and I've not used the HoNOSCA, so when I looked at it at the end, there's no sort of -overall cut-off score (2nd person again) -yeah, or something like that; I know we enter the points as separate items, but it doesn't give me an overall picture of it all
P4 – Key Principles	Education and Training	Sources of Information	The other thing I, in terms of, I don't know if this is appropriate or not, but, I don't know about you guys, but when I looked at it, I thought: *How much*

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
	or instruction		*of this am I meant to observe? How much of this am I meant to rely on parent reports?* And there's elements of it I think I'd actually have to go and observe the infant with other caregivers or in other environments, or talk to other environment care providers to see whether the behaviours are more global, all happening in that moment…have you already said this?[referring to participant having stepped out of the room to take a call], no, so I don't know, when I looked at the tool, I thought yes I could use it, I could see it, I could use it but how much of this is supposed to be parent self-report because I think the preamble talks about getting the information from as many different sources as you need (?),so (other respondent reads: *Use all available information in your rating to reflect your judgement based on all sources of information available to you*) [back to first respondent]…so that might be parent report, it might be care provider report, it might be what you observe, there's a whole lot of things and to actually answer some of these questions in a manner that I would think was actually accurate, I would probably do all of the above and so I'd be observing, I'd be asking, I'd be asking other care providers
P4 – Key Principles	Education and Training	Familiarity	And I suppose that does happen when you do, I mean I am only familiar with the adult one, but you do take information from all different sources, not only one, and when you do it regularly, you are sort of familiar, by habit -but there's the expertise of the person you're talking to, to be able to make the age-appropriate comments, so I might be able to make it while I'm observing it, or I might be able to ask the parent, but yeah, interesting, I don't know, I don't know whether that's harder, but that was my thought about it
P4 – Key Principles	Global	Length	I'd be interested to know how long, people who put this tool together, thought it would take, to be completed?
P4 – Key Principles	Global	Length	Its very lengthy to read, but I actually found that really helpful, um, probably *because* I don't think I'm an expert in infant development and stuff, I just thought that it gave a lot of really good examples and that it helped me to complete it really well
P4 – Key Principles	How to complete	Sources of Information	I think your point about, and maybe this is more relevant to specific scales, but your point about the parents' view of the issue versus the clinician view, versus the view of any, you know those things

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			sometimes are difficult to reconcile and so, um, yeah, kind of assimilating that information and having confidence in your judgement about rating
P4 – Key Principles	How to complete	Sources of Information	(another person:) and I'm just imagining trying to observe all of the circumstances that I'm being asked to rate
P4 – Key Principles	How to complete	Sources of Information	(another person) Well, I mean for example, Sleep. I mean you're not going to be sitting with a parent for example at 7 o'clock at night when they're trying to, you're relying on them to say this baby is impossible to get to sleep or you're relying in their expectation, so if their expectation is that their baby will sleep through the night, where actually developmentally that's not appropriate, you know, you need to have the skills to explore that with the parent in terms of *Is their expectation developmentally appropriate, or not?* Is there a mismatch there? Which then perhaps comes into the other scale around
P4 – Key Principles	How to complete	Which scale to use	...one thing I noticed that it said the scales should be rated in a certain order? And I'm interested in that's how the HoNOS and the HoNOSCA should also be done, so that should make it more uniform with those tools, OK. 'cause when I first went to look at it, I noticed the first one said now don't rate x,y and z in this scale, you rate that in that scale and I'm thinking, oh OK, (other person) and it does that in HoNOS as well (first person again) Yeah, yep, that's fine, so that was just an observation, I thought oh, um, but I understand them wanting to make it more uniform with the previous tools.
P4 – Key Principles	Education and training Structure	How to complete; order of scales Change order to scales Familiarity or instruction needed Why in order?	I agree. I had the same experience when I did a trial on a person, like in practice, but yeah, understanding the reason behind the instruction to rate one after the other and go from 1 to 15, if there's logic behind it, but I went through it and went oh, well, actually, and then I changed it around and I think that reflected better, I'm guessing was better Facilitator: -So you went though it once completely and then went back and changed some? Respondent: Yeah, 'cause I understood the whole measure, like I'm sure once you're really familiar with it, it would be different, but on the first time when I was doing my example kiddie, that's, like I had to go back; I don't know if the instruction is

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			really that helpful if there is not like a logic to why you have to do it in order. Other respondent: Why in order? and I wondered if it was about the uniformity of the tool, so making it more consistent with the other tools in the family of tools. Other respondent: And it may be in the HoNOSCA research that there is a reason given around that and why you do it that way I don't know – we've just always done it that way How compliant of you (laughing) I'm a good employee, aren't I?
Key Principles	How to complete	Difficult to follow instruction	I found it tricky to follow the Key Principle of *Do not include information already rated in an earlier scale*. [Agreement heard on recording] Because the scales are so often reflecting similar constructs, so I don't, yeah, I think its tricky from an assessment perspective, what's, for instance like especially the Scale 5 around *Developmental delays* and then we reference development throughout the entire thing, like its integrated across nearly every scale apart from the-how to complete the other ones which are around family, so like *developmental issues* could be related to Scale 6 and *Problems of physical illness* could be related to scale 7, *Sensory Processing Information* could be related to Sleep and other developmental stages, so following that instruction was tricky for me to not relate them (other respondent:) or to not use information from one rating or to have them impact on what you're thinking about the current rating, so even if you were to try and separate it out, it would actually be quite difficult to separate it out because of the integrated nature of all of those signs together
	How to complete	How to rate; not source of problem	I thought it was good the qualifier and you mention it also in the preamble I think, that, the presence of the clinically significant rating does not imply the source of the problem or the position for intervention – I think that's really nice, very relevant

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
	Global score	Global Score	(other respondent:) and that's why I think it is, and that's what would be helpful for me, if I was rating to understand, how the overall, what the overall rating means at the end, so I guess that its, I understand that its saying it's a measurement of the infant's mental health, I'm assuming… [Facilitator confirms: yes; social and emotional well-being] (Respondent:)…that's right, so its not an indicator for intervention (other respondent:) although they do say for all of these scales that a rating of 2 and above is clinically significant, which to me implies a need for some sort of intervention. If something's clinically significant then you can't rate it a 2 and then go Oh don't worry, you know, like, I think there's that sense of (other respondent:) but at the same time there's that sense of…it may not be the locus of intervention, with clinically significant it may not be what's driving the (back to previous respondent:) the intervention may not be with the infant the intervention may be what's driving it I think that's an important distinction [agreement heard] because otherwise you'd start addressing, say you start addressing all these say Scale 5,6, and 7 when the issue is really related to a bigger picture and this scale should not indicate what to do
	Purpose	How informs assessment process	So I guess then too, in saying that, then what might be useful in terms of the tool going forward is some statement around how they would see this tool informing the assessment process within a suite of tools to help make that statement that its not necessarily identifying a locus or a point for intervention, its actually identifying something that's happening with the infant or having an impact on the infant's social and emotional well-being, but I think then it might be useful, particularly if its going to be really broadly available and used for people and what we know if it comes in, and its used like the HoNOSCA and HoNOS are used, then there's gonna be non-infant-mental-health trained people picking it up and expected to use it
	Useful in practice	Usefulness	This would be really helpful, like when I was talking to *Brighter Futures* this week, to use the language to be really specific about what the issues were with this child, so in other services as well

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			[agreement heard] And like the adult ones used for NGOs as well, its a good way to discuss specific issues
Scale 1	Wording	Clarity	I had a particular issue with the word *settling* in here because it deliberately excludes sleep and I think that lots of people, with infants, when they talk about settling, what they talk about is settling them to sleep, which is actually addressed in a different scale, so you may want to look at another way of expressing because what I think they mean is being able to calm down and you know, regulate, um, rather than rocking them until they go to sleep, which is the commonly used way that parents will talk about settling, and that's quite specifically addressed somewhere else, whereas I would think about a different way of expressing that -calming down or -yeah, absolutely -yeah, just those specific little… -and when you read number 8 it says settling as one of the includes
Scale 1	Education and Training or instruction	Specialist Knowledge	I found difficult to separate what I *know* when I was reading this, it was a little but difficult, but I could see, yeah OK, I get that, but it was difficult to know what it might be like for someone who *doesn't* get it, who then has to actually use the tool – a bit hard to un-know what you know and it's a bit hard to try and look at it from the perspective of somebody who might be trying to be using it and they're not trained
Scale 1and Scale 10	Description	Clarity	On that paragraph where we're going more into specific examples, *Include difficultly*, I can see I think there appropriate specific example *undue irritability, excessive crying, frequently arching back*, maybe not so much in the sense of now you're getting really specific, whereas with excessive crying, undue irritability, people can imagine what it looks like, but anyway, going down to *inhibition of expected interactive response*, I found that one really vague and naturally linking more to other scales like the *social reciprocity* in scale 10, I am not sure what is meant by *inhibition of expected interactive response* so maybe they're trying to talk about not that the child doesn't sooth when we're expecting it to sooth, that's the *expected response*, but *interactive response*, is more we'd be expecting generally when we talk about interactive response, at least for me would be like the *social reciprocity*, more like the shared

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			eye contact for infants, yes for their emotion regulation to settle, but they do that in a social way (other respondent:) what they're presumably meaning is that when an infant's attempted to be settled, and they're really not engaging with that, I mean that's kind of what its looking at (other respondent:) I thought something different. I thought its about the situation where the infant's actually inhibiting their response to an interaction with the caregiver rather than just naturally responding (other respondent:) then that would be in *social* [agreement heard re interpretation of meaning] (other respondent:) what she said! [ie indicating agreement] (other respondent:) I think that makes sense, but not in the context of this scale…for what they might have meant in this scale (other respondent:)that would almost be *over*controlled emotional regulation and that's why it doesn't seem to fit, but maybe that's not what they mean …I mean it doesn't actually ask you to rate that under emotion (other respondent:)but it does talk about *psychological indicators of stress (hiccups, yawns, non-injurious scratching)* (other respondent:) yeah. I really like that part
Scale 1	Title	Scale	(other respondent:) but that is actually inhibiting, that's actually *overcontrolling or inhibiting* rather than under controlling so maybe the name of the scale could be reconsidered [agreement heard]: because then that would fit as inhibition would fit as if it was difficulty accessing help in managing their emotions (other respondent:) or they're not safe, they don't feel safe…and this is where I need to come back to think about what I am rating 'cause I don't want to start getting into interaction type stuff where you see, where you're rating the quality of the relationship and that's what I keep reminding myself of in doing this tools is I'm actually not, in most of the scales, rating the quality of the relationship
Key Principles	Global	Global Score	…that's what I'm talking about; it would be useful to talk about what other tools might assist and

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			overall assessment like rating the quality of the relationship, parental mental health, you know, that sort of stuff as well
Scale 1 and 9	How to complete	How to rate	They actually do talk in Scale 9 about that sort of restriction and affect and withdrawal
Scale 1	Description	Clarity	Need more clarity about disruptive behaviour
Scale 1 and 9	Description	Clarity	I find it a little bit like *Problems with disruptive behaviour /under controlled emotional regulation* I'm not clear what the topic of this scale is. Is it difficulties with emotional regulation, such as that they're expressed such as in disruptive behaviour, 'cause emotional regulation you could have the internalising kid, or the externalising kid and it kind of gives maybe overemphasis to disruptive behaviour which we already know that disruptive behaviour is disruptive behaviours that are gonna me more easily identified than the kids who internalise but they're still having problems with emotional regulation, so maybe the scale is meant to be emphasising that style of emotion dysregulation, but that's not like you're not seeing four month olds displaying disruptive behaviour for instance but you will see 1 month olds with poor emotion regulation and a tendancy to internalise or something, you would see that, so it might, you'll miss those ones (other respondent) well no you'll get them on scale 9, but I am wondering why we need, why those two things are being pulled apart 'cause Scale 9 is about emotion and related symptoms, so is Scale 1 really about behaviour?
Scale 1	Title	Clarity	I am struggling to figure out what it is asking for, to measure…it seems like its saying lots of things, trying to put into one thing
Scale 1 and 9	How to complete	How to rate	Scale 9 in the exclude box it says *Disruptive behaviours resulting from emotional distress – rated at Scale 1,* but the *emotion associated with the disruptive behaviour is rated at scale 9* – I don't know if that helps (other respondent) yeah that makes that clearer, but I'm not sure that it clears up, and because in fact they don't put exclude, like on Scale 1 maybe they need to talk about what you'd rate at Scale 9
Scale 1	Title	Clarity	Or change Scale 1 title to reflect if its about behaviour then *Behavioural problems following*

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			issues with emotion regulation instead of problems with both emotions and behaviour
Scale 1	Title	How to rate	Because it feelings like that box on Scale 9, the exclude box, that this is expecting you to rate behaviour and this is expecting you to rate emotion, but I don't think that's necessarily clear in Scale 1
Scale 2	Education and Training or Instruction	How to rate	There's a statement that says where two factors appear to negate each other, rate the most severe occurrence, I found that unclear and difficult to understand what I was meant to do
Scale 2	Description	Clarity	Rating 3 – *appropriate supports* - I am not sure what is meant by appropriate supports; is this referring to parental supports, or what sort of supports are appropriate?
Scale 2	Description	Clarity	I personally find *appropriate* a difficult word because that can mean different things
Scale 2	Rating	Interpretation	Rating 4 says there are severe levels of overactivity/underactivity which impacts negatively on the infant's capacity to engage and achieve developmental milestones across *multiple contexts*. What does that mean? Are we talking about caregiving contexts or life contexts or are we talking about multiple contexts of the infant? I think I went to environmental contexts.
		How to rate	I just said this could be very difficult to rate outside the caregiver context and then may require the assessor to try to interact with the infant, which, I do know that you guys, but I tend to think of not directly interacting with the infant but to observe the interaction…not sure whether this is now asking me to interact with the infant to see if I would get a different response from the infant to what I observe with their caregiver, so what I've said is this could be very difficult to rate outside of the caregiver context and then may require the assessor to try to interact with the infant to assess, that is assess with the caregiver and then assessment the clinician providing activity and appropriate support to see if the infant does things differently
Scale 2	Education and Training or Instructions	How to rate	I took the multiple contexts to mean that it's more pervasive, so you might from the day care giver, is not limited to one personal situation
Scale 2	Education and	How to rate	So that means you've either got to go on watch it, or you've got to talk to people

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
	Training or Instructions		
Scale 2	Education and Training or Instructions	How to rate	So do we go with the majority then? … What I mean is the tension side of things usually becomes an issue with one particular person, usually the mother, and it may not come across in other people, so what are we measuring?
Scale 2	Education and Training or Instructions	How to rate	You're measuring the infant. So in that case you wouldn't rate 4; it's not sustained across *multiple contexts,* you'd rate 2 or 3, that would be the difference
Scale 2	Wording	Clarity	I think that there is a little bit of repetition in the description *associated with any cause, including…* then in the next paragraph you say again: *due to any cause including depression*… Instead you could use some text to outline some examples of infant difficulties with joint or sustained attention, what that would look like, 'cause if a person is not used to how long should a baby be looking at Mum for, is that normalising, I think some breaks in eye contact but actually saying there is a noticeable problem with the child
Scale 2	Education and Training or Instructions	Reference to developmental stages	And this is with a talk about it going to be in age development, because I've actually seen earlier versions of this document and one of the issues will be what a 3 month-er can do, is going to be very different to what a 47 month-er can do, so therefore the person doing it needs to understand the difference between the two and I'm not clear that will happen
Scale 2	Education and Training or instruction	Reference to developmental stages	I think something needs to go with it. There needs to be some training, and it needs to be some reference to the different stages
Scale 2	Education and Training or Instruction	Reference to developmental stages	So maybe some recommendations about who is qualified, or who is benchmarked, whatever term they want to use, to know the difference between using the tool on the 47-month old and using the tool on a three-month-old
Scale 2	Education and Training	Reference to developmental stages	That's probably not that dissimilar to HoNOSCA in a lot of ways because you asked to rate 4-17, and if you don't have a sense of what those developmental stages should look like…
Scale 3	Wording	Clarity	This one is tricky saying *self injury* because I'm excluding baby infant-infants straight away... and then as I read further, I thought oh, its also a little bit of that self-soothing, injury is joint under this box, maybe it's just semantics, but I usually went "Oh! That is never applied to my under-1's" and

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			then as a read more I thought "Oh, under this scale it would, um [unsure sound]"… probably just that word *self injury*
Scale 3	Wording	Clarity	And then it kind of relates to *intention*, which I think you guys are trying to flesh out in those lines, you've been talking about intentionality with a baby it's pretty hard anyway
Scale 3	Education and Training or instruction	Reference to developmental stages	And the transition from the baby to the one that deliberately harms himself, maybe my two-year-old who is deliberately being himself on the head with something doesn't know what that's about but that can escalate into an older, deliberate thing, it's not clear, again, about the stages, what sort of happens at each stage
Scale 3	Clarity	Clarity: Interpretation	*Lack of self protective reflexes* - is that kind of a different thing to the child performing self injury, by not protecting myself, it also means I'm self injuring myself? If I was not to protect myself as an adult, I could get that that was about self injury
Scale 3	Title	Clarity: Interpretation	So even going to *non-accidental self injury/lack of self protective behaviours* as a name in the scale [Agreement heard] I think that's good 'cause I think that name [referring to existing] will trip people up… (Other respondent:) I was thinking you are more trying to get at dysfunctional self soothing behaviour rather than the child trying to intentionally harm itself, but maybe I'm misunderstanding it Facilitator: and do you mean that across the range, or for the younger Respondent: Across the range, because even a 4 year old who might hit themselves on the head, in some ways its trying to get their needs met… It's a way of self soothing that's not working, so I suppose that comes from the way you see people's behaviour, if you see the behaviour is always trying to get their needs met
Scale 3	Rating	How to rate	The medical condition, does that include autism, that sort of thing? (Other respondent:) and should it give examples if it does… (Other respondent:) …you'd be rating it anyway… because it is dysfunctional (Other respondent:)…but when you have the condition, your capacity is on a different level (Other respondent:)…but then you're accounting for condition for everyone

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			(Other respondent:)… That's what I don't understand in terms of this is a scale that measures the capacity, is that correct? (Other respondent:)… Does exclude secondary to medical condition but then on scale 6, it then doesn't talk about self injurious behaviour so theoretically you could have a three-year-old headbanging, that you would rate zero…because it was due to for example being on the spectrum but then you wouldn't rate that anywhere else, if you're excluding it here…if you're excluding it here, it probably needs to come in somewhere else
Scale 3 and 1	Rating	How to rate	… On Scale 3 is says exclude *Self-injurious behaviour secondary to a medical condition* but in Scale 1, I had to *include behaviour associated with any disorder* and so I wasn't sure why I was including it in scale 1 [behaviours, whether disruptive or self-injurious] and excluding it in the third Scale…and maybe I don't need to know – maybe I just need to do what I'm told (laughs)
Scale 3	Rating	How to rate	But it also does instruct you in the instructions not to consider the causes, so it is a little bit tricky – I understand when they are saying exclude because you're rating on another scale, that sits perfectly comfortably, but just saying *don't worry about it*, like I wouldn't feel comfortable rating a little kid headbanging on a zero just because he had a diagnosis of ASD
Scale 4	Rating	How to rate	To accurately assess this, I'd need to observe feeding, because so many of these things, if they're parent reported are much harder to rate for accuracy when the parent's reporting it, so I would need to observe it
Scale 4	Length	Length	so that then goes to my previous question about thinking how long it would take me to use this tool
Scale 4	Rating	How to rate	If you're being objective and saying someone recognises that this is a problem, even if it is just the mother, and you're rating it based in that, then it is sort of flagging something that needs to be discussed with someone who can better assess it, which I wouldn't be doing for feeding behaviour, but you know quite often they have someone like an early childhood nurse or someone that can assess that, and I suppose if it is a tool to have a conversation with people, to identify problem areas, then its better to overrate than underrate, isn't it, even if it is just the mother that's concerned about it. Next time you rate it at 13 weeks, or

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			whatever it is it might have been fully resolved and then that would reflect the new rating
Scale 4	Wording	Clarity	What does *Mild but definite problem with non-organic somatic symptoms* mean? [Rating 2]…because if its not clear to me, it might not be clear to someone else
Scale 4	Wording	Clarity	*Moderately severe* – is it meant to be moderate?
Scale 5	Wording	Clarity	There was one that said *mildly severe*, which I thought was a really odd thing to say, can't remember which scale it was in [Scale 5]
Scale 4	Education and Training or instruction	How to rate	My only thing with this and I don't know how you capture it, is that feeding, like sleep, the developmental stage is difficult, so if you are measuring this during the infant's transition to solids, say, then they could rate much higher, you know then you could have parents who are struggling with that…I don't have a solution but I think maybe there needs to be a bit of a recognition about rating that occur during times of transition (other respondent:)Maybe even in a general comment right at the beginning (first respondent:) yeah, like if the child is starting daycare while there's a major life transition, it's probably not the right time, or they're simply tired (other respondent:)You watch a baby feeding when its tired compared to a baby feeding when its awake and playful… absolutely there could be quite a difference - it comes back to how that tool is used by the rater, do they know, are they aware enough to say is this how it usually is? to the parent
Scale 4	Education and Training or instruction	How to rate	I would be reluctant not to do a higher rating because they're at a transition phase though, because clients come in contact with us at transition phases because they are having problems, so excluding the fact they're in a transition would not be helpful because you're going to mix (other respondent:) no I didn't necessarily mean excluding, I just mean that… (other respondent:) sort of regardless of whether they're in a transitional phase or not, because the transitional phase will resolve itself next time you do the measure again
Scale 5	Education and Training or Instruction	Sources of information	I came back to the same thing about [rating 2] *mildly severe concerns noted across more than one setting and in comparison to peers of a similar age* – not an easy "do" ie not easy to collect that

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			information. It's not just something you can collect on observation; you'd have to be talking to multiple care providers or settings.
Scale 5	Wording	Clarity	[Rating 2] *mildly severe*
Scale 5	Title	Clarity	Referring to the younger population; communication skills: on this scale are you trying to assess communication skills or just a verbal communication skills? [Participant gives example of a three-month-old and makes a noise and contrasts this against eye contact]. [The suggestion is that "communication skills" may be too broad if we talking about verbal communication skills. Distinguish between global and verbal communication skills].
Scale 5	Description	Clarity	There are not many examples on the younger age
Scale 6	Description	Clarity	Does not mention self injury
Scale 6	Education and training or instruction	Sources of information Clarity	[this scale] brings in the reports from others. Because this scale is overt in stating that in the inclusion, it made me think "does that mean I shouldn't include information from others in the other scales".
Scale 6	Wording	Clarity	The Scale references self injurious behaviour twice; double up
Scale 7	Education and training or instruction	How to rate	This scale felt very difficult to rate. It may require a series of specific interactions to be able to observe and scale it, and how much would the scale respond to a parent being asked questions?
Scale 7	Description	How to rate	Good that it's fleshed out as it may not be in the knowledge scope of a lot of people who'd be rating this, and it is important.
Scale 8	Wording	Clarity	[Rating 3] *The infant's sleeping pattern is a cause for distress in the parents and family.* Depending on parent - Parent can't sleep, or can't cope. Seems unfair to rate the child according to the parent. It may only matter that the infant has a sleep issue; it may not matter what the cause of the issue is. … Parent may have a low tolerance. Is the scale is supposed to pick that up or not?
Scale 8	Rating	Measure	There is a big jump between 2 and 3 on this scale and not much of a gap between 3 and 4.
Scale 8	Description	Clarity	Give more examples in the description like insufficient sleep for younger infants. Most of the examples in brackets are for older babies. For example "difficulty resettling" for junior babies.
Scale 9	Scales	How to rate	The exclusion states "*physical sequelae of psychological disorders or medication – rated at*

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			scale 6". I thought Scale 6 was focusing on physical, not psychological.
Scale 1 and 9	Scales	How to rate	Behavioural versus emotional. Crossover between [Scales] 1 and 9.
Scale 9	Description	Clarity	Not focusing on junior [under one year] babies that much. … Harder for your average clinician to see…thinking of CYMHS using this tool. You can relate 2-year old behaviour to something but to do that with a week-old baby or month-old baby is a very difficult picture.
Scale 10	Education and training or instruction	Sources of information	*The scale includes problems with emotional attunement and misalignment between the infant's needs and the parents' or carers' responses.* The only way that could be measured… Is to be observed.
Scale 10	Description	How to rate	Need more examples around older babies [of social reciprocity]
Scale 11	Description	Clarity	*Self-care* and *exploration* are such different topics. … Self-care might be about the older baby. Exploration may be about the younger baby; Might be looking at the same thing, with the age differential? So maybe that needs to be clarified in the statement, if that's is correct.
Scale 11	Scale	Useful	Environmental exploration is a good topic to have as a scale; helpful.
Scale 12	Rating	Clarity	[Rating 1] *Parental reflective capacity* needs to be well defined. Do untrained people know what this is? Spell out more.
Scale 12	Rating	Clarity	Distil it further between rating 2 and rating 4. In rating 2 there is a problem in the relationship for the child, but there is another significant person like at pre-school or grandparent, whereas in rating 4 - there is a serious problem in the relationship plus there's no one else … No functional relationship
Scale 13	Rating	Clarity	If the parent is not bringing the child to playgroup or any other social setting it would be rated as a 0, but it should not be a 0. … They talk about the exclusion in scale 11, but maybe it needs to be included in 12 and 13. … Or maybe the parent not taking the child needs to be included in rating 1,2,3, or 4 [agreement] *Include refusal, include limited or minimal opportunity…* Rating is about the infant, not the mother. Eg. if the infant is not getting the opportunity to socialise [daycare, cousins, friends]

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 14	Rating	Clarity	Normalising – first-time mum with newborn infant won't know, so we don't want to give a zero [distinction between that and major lack of knowledge]
Scale 14	Rating	Clarity	[Rating 0]: *parents demonstrate a good level of understanding about the difficulties* suggest to add: "or where to go to get information".
Scale 14	Description	Clarity	With this scale I find it hard to separate those parents who kind of know that there is a problem, but don't know where to go [from rating that at Scale 15]… Some mums may have difficulty recognising there is a problem… I think exaggerating difficulties is important too.
Scale 15	Rating	Clarity	Rating 1 description matches a rating for 0.
Scale 15	Scale	Rating	Needs to include if they have the knowledge, and willingness to go, but do not follow through or can't get there.
Scale 15	Description	Rating	Needs to include that cohort of parents who know what they need to do, they know they need help, but they are so scared about what that might mean, they then don't. May have had a negative experience in the past… If I ring mental health, that means FACS.
Scale 15	Description	Rating	What about parents with religious beliefs – may impede on assessing services … There are lots of other reasons people would not access support; it's not just about lack of information out there.
Overarching question Do you feel the instrument reflects the prescribed objectives of the instrument?	Global	Useful	Definitely facilitates discussion of infants' and young childrens' presentation amongst clinicians
	Global	Specialist knowledge	I think it's got the potential to document the progress of infants and young children and make overt changes, but it comes back to the wide range of changes that occur between 0-47 months. I think that the tool has the capacity to pick that up, if the rater has the capacity to pick that up. My concern is

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			that the tool requires some statement around who's going to be able to use it the way it was intended, because if you look at the group who put it together– highly specialised clinical people, but then I wonder about someone else picking that up
	Global	Continuity	I would think this would be one of a suite of tools used by and discussed by a MDT
	Global	aims	I would not say you would do this and it would inform treatment – it should inform assessment. For the other two aims I would say: tick, tick.
	Global	continuity	I think it's great that there's is a tool that's being developed that will fit alongside other tools that are accepted and regularly used and clinicians look at ratings, that makes sense to them. They used to looking at five-point scales, used to knowing what that means.
	Global	Needed	Having an infant-specific outcome measure is great.
	Global	Length	Initially it was clunky and cumbersome, but you'd be doing that day in and day out really. … quite long, but once they get used to it…
	Global	Length	The glossary is very comprehensive, but it doesn't necessarily address all of the nuances of this age range which is *such*…developmentally variable
	Global	Needed	I really like it, I feel it fills a vacuum, actually.

Group 2

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Global	Rating period	Time frame	So the rating period's two weeks. I think that's a good framework because you need some of that time because some of these parents may think that they're worse than they are, or not as bad as they are in the rating period and because it's only two weeks that can mean that it's hard to measure really in a way that will necessarily indicate how much of a problem it really is
Global	Rating period	Wording	Related to the timeframe, I think its scale 13 about attendance, socialisation and stuff - there is a comment that says if it's in a holiday period you can use the last two weeks, I'm wondering then should all of the scales be rated on the same two weeks using that scale as opposed to using say, the last two weeks for all the other items and then a different two weeks for the…[pause] because I think the two would influence each other potentially. Having just seen a six-year-old boy with autism this morning, where it makes a big difference, that was one of the things I wondered about with that. I mean it may not be as big of a deal with younger kids, but it's just a thought I had about would that materially affect other scales if you were using two different datasets to rate it. (other respondent:) a couple of things. If you make an exception for one measure, that's going to be difficult for people to hold that in mind, because you do so many of these. The other thing that I'd be worried about is if we made the exception for that HoNOSCA scale, should we do that for all the other HoNOSCA scales that we do? And are they going to at some point be, the data compared the HONOSCAs for a six-year-old, a two-year-old and a 10-year-old. If you make an exception for one, you have to do it for the other (first respondent:) my sense would be… to just do the whole scale the same; if you're in the second week of the school holidays and you're gonna do it on the last 2 weeks of school, you just do the whole scale on the same period of time, and it was confusing then [agreement heard] ... But I'm sort of ambiguous in that, and I wondered what people thought of that? (other respondent:) two options: One, it's a research study, does it make a difference? If you do the two weeks, not two weeks, and there is the big issue that my heart's fluttering about, if you change

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			it for this, you actually change it for all of them, and [sigh] but its worth thinking about for future iterations of HoNOSCA (first respondent:) I mean is it easier to remove that ambiguous sentence?... In a sense we leave it at the discretion of the rater with the HONOSCA as well (other respondent:) which is what I think happens anyway, when you're doing it, we do for kids on school holidays, especially over January and stuff, and it kind of… you have knowledge of the family… sort of on the ground I think that's how it works. You just adapt it without sticking too strictly to the two weeks
Scale 1	Wording	Clarity	With the diagnoses: depression, autism, one of the difficulties I had [referring to a 24 month old with ASD], does that have a rating of two or above because he already had a diagnosis, but the diagnosis is quite mild and wasn't given by our service so there is also a bit of disagreement around it, so I found that a little bit tricky My sense of HONOSCA historically is that you try to put diagnoses aside and just think about broadly, problem areas. So it doesn't actually matter whether the child is depressed or has an intellectual disability on the spectrum or how bad is it. It's just isn't there, or is it not there and you leave the speculation for other stuff. That's how I've done the HoNOSCA. I don't know if the HoNOSCA scaling actually talks about diagnoses… made me a little bit confused, probably if it wasn't there I wouldn't have thought about it
Scale 1	Education and training or instruction	How to rate	With rating 3 and 4 it's important to consider the duration and I also wondered whether you excluded traumatic experience from the parents
Scale 1	Education and training or instruction	How to rate	For the rating points down the bottom, some have less explanation, some have more, so it needs to be a bit more consistent for me understanding and what would *moderate* mean versus a *mild*. The moderate has got the explanation
Scale 1	Additional tools	Assessing accurately	When parents are reporting this… There are tools… I was thinking about the duration… They can actually give them data to then be more specific around these scales because parents might think it's all the time, and in the record for a week, and it's actually three times a week for 10 minutes, but

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			we feel like its all the time, therefore it comes out as a moderate or severe... That's where frequency and duration comes into it - parent informer as well; parents with mental illness often catastrophise the rating [Agreement heard]
Scale 1	Glossary	Assessing accurately	(Other respondent:) that could go into the glossary is well, other things to include (Other respondent:) clinically significant tantrums versus normal tantrums (Other respondent:) he doesn't stop moving; he has ADHD!
Scale 1	Education and Training or Instruction	How to rate Wording	I was unclear if I was going to rate in an increment, what you mean by disruptive behaviour in an infant. I don't know if that word is appropriate; an infant who's quite passive and listless the behaviours' of concern but its not necessarily disruptive… where you know they can't receive any care from their environment, where does that get coded, because that's quite significant? (other) overcontrolled emotion regulation …Yes, exactly. So where does that come in? …Scale 9
Scale 1 and 9	Scale 1 and 9	How to rate	There do seem some crossovers between 1 and 9
Global	Glossary	How to rate	That may be something you put in the glossary, for example overcontrolled regulation be put in Scale 9.
Scale 2	Scales	Structure	Why put activity and attention together? You may have problems in one or the other domain and clinically if you have both, it's much more serious. I wonder if we should have two scales for those? I'm thinking about trauma and developmental disorders. I am thinking of kid who's got a very high activity level and very poor attention as opposed to someone who's got an active temperament but can sustain attention.
Scale 2	Description	Clarity	This might be where listlessness might be noted as well; gives a different emphasis.
Scale 2	Rating	How to rate	If you have two problems of moderately severe intensity, is that a reason to score it as a 4? Rather than 10/10 on one thing, can two 7's [mean he same thing]…making it explicit rather than implicit – I think people will probably do that anyway.
Scale 3	Description	Clarity	Re biting – I have seen kids biting to self soothe themselves but they don't necessarily bite others to self-soothe –that was a question as to what that second statement was about?

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 3	Description	Clarity	The term self harm in infants bothers me. The term self harm in adolescents has a whole different connotation [suggestion: use another word – self-injury?] … The connotation is slightly different. We don't tend to refer to adolescents as self injuring. …be consistent with title.
Scale 3	Description	Clarity	*Non-intentional accident* needs to be defined
Scale 3	Description	Clarity	*self protective reflexes* needs to be defined
Scale 3	Description	Clarity	*Lack of safety awareness* eg jump out of a tree because they're not judging the distance or thinking of the consequences [include]. Non-intentional accident.
Global	Scales	Where to rate	*Self soothing behaviour* and some of the sensory seeking behaviour in another scale – double up …[Suggestion] where you code sensory seeking, refer back to where you code in Scale 3.
Scale 4	Scale	General	Straight-forward.
Scale 4	Description	Clarity	Description about what to include is quite clear
Scale 4	Description	Clarity	Thinking about those early days of breastfeeding versus bottle, the parent can actually be quite defensive as some hospitals promote breast feeding and some don't, and then there's a whole movement outside of CYMHS of breastfeeding versus bottle and that may affect how parents may look at feeding in those really young infants. The scale is clear but the political overlay in that context is going to affect those parents and this question. … Parents don't answer the question. It's a subjective assessment by the clinician. The clinician will be able to assess regardless of the political connotation.
Scale 5	Description	Clarity	What do we do with kids who have a global delay?
Scale 5	Title	Clarity	Put motor in the title … The first time I read it, I thought it was just on language and communication. … Helpful to have a paragraph or a sentence detailing what the motor skills are. ? Fine, ? gross ?Coordination ?balance
Scale 5	Global	Scales	Some issues are too broad and general, others too specific.
Scales	Global	Scales	Some scales seem to cover a lot of symptoms; can be confusing as to what is being looked at - perhaps break into dot points to make them more understandable
Scales	General	Clarity	Language ambiguous at times – be more simple – use plain language eg putative.

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 5	Title	Clarity	Rename: Problems with communication, motor or specific developmental delays - because language, I would argue, would fall under the rubric of communication - expand on what to include/exclude
Scale 5	Description	Where to rate	Motor planning could be covered under activity?
Scale 5	Education and training or instruction	Specialist knowledge	I looked at this and thought I need a speech therapist in a way, to have a really good look at this area… People may think they've got a language problem when they've got an articulation problem [motor]
Scale 5	Description	Clarity	Re: exclusion of intellectual disability. Hard to determine – suggestion – take that sentence out
Scale 5	Description	Clarity	It's good how you've got the include and exclude… Needs tweaking, but I like it.
Scale 6	Scale	General	Straightforward
Scale 6	Rating	Rating	Rating Scale is good; quite descriptive
Scale 6	Description	Clarity	Exclude: processing limiting physical activity, or coded in other item?
Scale 7	Description	Clarity	Wordy, particularly with examples at the end
Scale 7	Description	Clarity	Our team thought there was great detail about the sensory components, but perhaps *abnormal motor tone, floppiness and frozen watchfulness* are such broad issues and may not reflect sensory processing abnormalities, it may be better not to put that in.
Scale 7	Description	Clarity	Err on the side of being more prescriptive …there or in glossary
Scales	General	Consistency	Wondered if each scale is written by someone different; need to make it more consistent
Scale 8	Description	Clarity	Good - clarified that it could be with the parent or the environment
Scale 8	Description	Clarity	Not sure what that first sentence is supposed to imply – normalising.
Scale 8	Description	Clarity	*Disturbance* – change to *difficulties*
Scale 8	Description	Clarity	[Rating 3] *The infant's sleeping pattern is a cause for distress in the parents and family.* Will depend on age; if older and still highly disturbed, its problematic. Put more explanation around that. [Suggestion:] "Sleep patterns can vary significantly with age."
Scale 9	Scale	General	Trying to encompass an extraordinarily broad range of things – half of the DSM in one scale
Scale 1 and 9	Title	Clarity	Internalising – withdrawal, listlessness – reflect in the title, as opposed Scale 1. [Suggest:] internalising
Scale 9	Description	Clarity	I think clinicians will be able to rate that. You don't need to be overly inclusive. Its something that we're pretty familiar with.
Scale 10	Scale	General	I think it's a really good Scale; I'm glad it's included.

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 10	Rating	Clarity	There are two jumps that happen within that scale: you go from [Rating 2] *Mild but definite problems in at least one area* to [Rating 3] *Moderate problems in more than one area.* Its like its sort of fused - you go from mild, moderate, severe, and how many areas, sort of fused into one. … Or just drop the whole thing. Have mild problems, moderate problems, severe problems.
Scale 11	Education and Training or instruction	How to rate	Do you consult with other people to complete it, or is it purely based on the clinician's observations? …Reflect caregiving level and the parents ability to recognise the child's development capability, so you do need observational data in this setting. … It's a very subjective rating. If its consistent over time, that's the important part.
Scales 11,12,13	Scales	Where to rate	In general, I think there is a lot of cross-pollination over the next three items [Scales]; how environmental exploration interplays with family life, interplays with opportunities for socialisation, and to some extent I think it will be sort of artificial how we pick those things
Scales	Education and training or instruction	Specialist knowledge	I think there's learning effect – the more you do these, the more it sorts itself out; where to rate this, or where to rate that.
Scales	General	Continuity	I really liked the fact that you actually looked at the context of the HoNOSCA.eg, they've got 4 other kids under five, the husband's violent, whatever it may be… even including it, whatever the content is, for me, it felt like a breakthrough, and I was very excited, so I loved it; really, really good, particularly at this age group…because it brings in that issue… about the trauma [its usually] only the dyad work, and that doesn't reflect the whole context
Scale 13	Scale	General	Straightforward
Scale 12	Description	Clarity	No use of the word *trauma* in there.
Scale 13	Length	Length	On the verge of what would be considered long in the include box … The shorter, more succinct, the more practical it will be in day-to-day use
Scale 14	Scale	General	I think it just forces people to ask a question that some of us routinely ask anyway "tell me what you think's going on"
Scales 14 and 15	Education and training or instructions	Specialist knowledge	My only comment is that service providers are often confused about the system and knowledge – it's almost a given that there be problems about knowledge. Translating that into accessing the system is also different – because people may have knowledge, but they may not; I thought that [Scale]

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			15 was really good – managing how to ask for assistance and also navigate them are a different skill
Overarching question	Global	General	I think it does [agreement] … Certainly does it better than what we had before, which was nothing! … We've had to use HoNOSCAs that were nonsensical in parts, for infants

Group 3

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Global	p3 – Index of Scales	Index	Similar to the HoNOS and HoNOSCA - I think that's probably good in a way, it's familiar
Scale 1	Scale	General	Got our tick of approval
Scale 1	Wording	Clarity Specialist Knowledge	Title – the words *disruptive under control emotional regulation* - I wonder possibly about the word *dysregulated* might be more appropriate for infant clinicians, but …not as clear for people who aren't; haven't worked in that field. The word disruptive or challenging, they tend to me, to be developmentally appropriate for older kids and I wonder whether it could be added
Scale 1	Scale Structure	Structure Title	We thought that this is something that perhaps should be separated, the disruptive behaviour from dysregulation… Certainly someone could be dysregulated, and nothing to do with disruptive … Yeah, and that would come up in other scales
Scale 1	Ratings	How to rate Title	The severity ratings then, it's become synonymous with aggression … Would it be a better way of titling it as an aggression scale … Yeah, because dysregulation is under sleep and feeding and other sorts of scales … Then how would we describe aggression in a six month old … Make it clear that this scale is about aggression… I think the word disruptive is…Is it an aggression scale? If it is then we should call it an aggression scale.
Scale 1	Clarify	Ratings	Might want to include toddler defiance which isn't in that group …excessive tantruming ... Aggressive and defiant behaviours, which is sort of alluded to in the text but not so much in the descriptors down the bottom [ratings]… You might put some defiance to request by parents but is not pervasive and parent can convince the child to do things at some of the time, can overcome these with persistence
Scale 1	Scale	How to rate	Is it a behaviour scale? … I still think it should be separate [a separate scale – Disruptive behaviour separate scale to dysregulation]
Scale 1 and 9	Scale	How to rate	Dysregulation should probably come under Scale nine… So just extract the dysregulation symptoms... and put them into the examples for the emotional.The dysregulation is described in the

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			descriptor in a non-aggressive way, and then also in an aggressive way
Scale 1	Education and Training or instruction	How to rate	What do we rate it on? I assume it's over the last two weeks, so parental descriptions will have a large weighting, whether you rated differently on the mental state observations and kind of like what you see
Scale 2	Scales	Structure	Again we were wondering about two scales here; the problem with activity levels, separate from joint and/or sustained attention ... Yeah, yeah, that really stands out doesn't it? There so much to take in.
Global	Length	Length	We're thinking to summarise it a bit, generally, if there could be less words…
Global	Structure	Consistency	It looked like there might have been different people describing it and the language changed, so if there could be one editor using the same language all the way through … I agree. I think as we move into the instrument, the severity ratings become less wordy and I think it's kind of clearer that the clinician makes a judgement about mild, moderate or severe, rather than trying to find examples, because people get caught up on the examples, I think. Clinicians do when they're rating it, they think "oh no, I don't have that particular thing", but they maybe missed some of the different… What I like about it, one of the improvements from the previous versions we've tried, was that this a bit more in younger infants and preschoolers,… Divide… Which I think is helpful for clinicians, particularly if they're not as confident with younger infants, to have some examples of how things might present in an infant. I think the examples of there, it might just be about structuring them in a consistent way, for younger infants and older. [Generalised comments to put examples in descriptor, not rating]
		Specialist knowledge	
		Developmental stages	
Scale 2	Wording	Clarity	Swap around the sentences on rating two, it just flows with joint attention is the second point in all of them, except in number to where it's the first point that's made
Scale 2	Wording	Clarity	Joint and/or sustained – make consistent
Scale 3	Scale	Structure	It feels a bit like a forced scale in some ways, its sort of adapted obviously from the HoNOSCA. You do see occasions of self injury, probably more in toddlers, probably less scalable I think than this

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 3	Wording	Clarity; bad example	We were a bit concerned about having them making dolls hit their heads during play; that can happen with a kid who's got nothing outside, to be concerned about, not self injury, just part of play ... It feels like it should be almost a more binary type of rating in a way rather than trying to kind of tease it out into different components. If you have kids who are headbanging or slamming their heads into a wall or something, I think that is worth rating, but is it possible to actually tease out the little bits in between, it feels a bit forced - to fit a model rather than a scenario … We wouldn't be thinking about self injury at all for kids who are banging their dolls together
Scale 3	Description	Clarity	Lack of self protective reflexes… Would be used under unusual circumstances
Scale 4	General	Scale	Important scale to have – see a lot of …absolutely
Scale 4	Description	Rating	Rating descriptors could be tweaked a bit…additional support to feed – could be interpreted with naso-gastric feeds…dump those examples and let the clinical decide
Scale 4	Scale	How to rate	Does it include the relational aspect – rate rather than exclude relational
Scale 4	Description	Clarity	Include: independent eating for older child
Scale 4	Description	Clarity	Might get a kid who is feeding OK [nutrition and observed] and is all right, but you get a parent is stressed out of their head – we would you rate parent anxiety about feeding in their mind? Rate as 4 or 1 or 0?
Scale 4	Wording	Clarity	… You make it then problems with feeding, rather than problems with feeding behaviour?
Scale 4	Description	Clarity	Add in descriptor parental concerns regarding feeding, otherwise it won't be captured if there's a perfectly well infant
Scale 4	Title	Clarity	Changed to "Problems with feeding"
Scale 5	Scale	Structure	Trying to capture all of those things … We have whole teams to work that out
Scale 5	Description	Clarity	[Related comments from speech pathologists who were not present, but had passed on their input:] Description does not include comprehension, sentence structure, vocabulary as a language or communication difficulty
Scale 5	Description	Clarity	Also seems weighted towards communication, even though it's meant to cover all of the developmental domains…
Scale 5	Description	Clarity	Have them separated to say language difficulties cognitive and problem-solving, which most difficult is include, it doesn't include toilet difficulties (in self-care)… There is quite a bit about language

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			because it is appropriate; the language is the most obvious which brings kids to the attention of services whereas motor and cognitive difficulties are more subtle, so we need more examples of those, just like with language, so people who aren't as clear, it will be clinically useful for someone screening cognitive and motor… add a sentence for each of those
Scale 5	Description	Clarity	Exclusion for global development delay is confusing. This in the development delay item, but then it excludes developmental delay; that's just cognitive version of a specific developmental delay. These distinctions become less meaningful as you move into infancy… Once you get to 6 or seven, it evens out… Take out the exclusion of developmental disability.
Scale 5	Education and training or instruction	How to rate	If you have a delayed infant, where is that recorded?
Scale 5	Title	Clarity	... Take out the exclusion. … or just take out the "specific" … And I'd say " language, cognitive and motor skills"
Scale 6	Scale	General	No specific comments to make … Looks all right
Scale 7	Description	Clarity	Take out "abnormal motor tone", would go into an example of a motor difficulty in the other scale, and isn't specifically related to sensory
Scale 7	Description	Clarity	Presumably you're excluding feeding sensitivities from here [suggested to say that in exclusions] would come under feeding
Scale 7	Description	Clarity	Sensory processing could go to sensory regulation
Scale 8	Education and training or instruction	Rating	With scale, but so subjective; you could get such a range of ratings – for example – could get a kid of six times a night and it's fine, or a kid up three times a night and it's a disaster, and that would be parent rated and also clinician rated. I think there's a lot of clinician prejudice in how they might rate this scale in particular.
Scale 8 and 4	Description	Clarity	Like that [it] accounts for parent distress in the family and I would like to borrow that in feeding as well; same perception issue in those two regulatory domains
Scale 8	Education and training or instruction	How to rate	Also begs the question where you got our family who's really distressed, or even one family member maybe a mother who's really distressed, dad's not distressed, clinicians saying well actually this kid

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			looks well rested, will actually its getting a reasonable amount of sleep; how do you rate that then – do you rated on what you're seeing with the baby, what you're seeing with mum you know tearing her hair out… Maybe you take these factors into account and coalesce them into some sort of mid range type rating?
Scales	Global	How to rate	The use of the scale in its entirety will have a lot more clinician prejudice than older kids
Scale 9	Description	Clarity	May include changes in curiosity, emotional responses, excessive stillness and restrictions in affect range, increased tolerance for aversive adult behaviour [make the inclusions dot points]
Scale 9	Scale	General	I actually like this one. I think its pretty good … Has different developmentally appropriate examples – I think that's good … Encapsulated well…Delineated examples
Scale 10	Description	Clarity	How is the infant engaging in early relationships or is this kind of an early sign for autism kind of scale? [Preference – not to be an autism scale]
Scale 10	Description	Clarity	Descriptors could be a little more specific. Look at it as more of a temporal concept, rather than one area: "The infant chose to engage in reciprocal social interaction at some time with some people", rather than in one area, or "the infant showed no capacity to engage interactions with anyone". A bit more specificity would help the rater.
Scale 11	Description	Clarity	Tricky one… Its adaption from all the children and adults… It's a clunky concept in my mind. … There's a few things – it's toileting [regulatory difficulties], adaptive behaviour [adaptive self-care involves feeding], which we've covered off elsewhere. Environmental exploration – you can get the trauma pattern… or the aloof disconnection… So clinically if you think about the clusters of symptoms you would see in this category, it draws on a lot of different possibilities ... The other thing is, from preschool, self-care is such a relational concept, this struggles to capture the relational element of it; can this person seek help from someone to get dressed? Can they signal their need to go to the toilet?… That's where it grates. It probably lacks that relational element, and are we measuring something else if we include that?… My worry is… If you get a kid who's really avoidance in their attachment and does everything fully independently, then that might rate really highly on this scale and yet its quite maladaptive

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			because they're not able to use relationships to get their needs met. … Over reliance or under reliance on caregivers for support
Scale 11	Description	Clarity	Have some early examples for little ones, particularly of exploration… In that earlier age range, you're looking for exploration, not self-care.
Scale 11	Title	Clarity	*Can they use caregivers to meet their needs?* That would be a better title [agreement heard]
Scale 12	Scale	General	I like this scale
Scale 12	Description	Clarity	I wouldn't mind seeing *parental reflective capacity* which is mentioned [Rating 1]… Either put in the more severe ratings, or remove from here and put in the descriptor; to make it clear that its practical physical resources as well is parental reflective capacity that are being rated on this scale
Scale 13	Education and training or instruction	How to rate	This is an interesting one for infants. This is the equivalence of school attendance/work [referring to scale being derived from other scales] … In infants, is that about separation with other caregivers? What symptoms are pathological in a three-month-old difficulty engaging in child care? I think that's the challenge; what the symptoms are.
Scales	Global	Description	The infant, the parent and the relationship – those three perspectives are relevant to all scales
Scale 13	Ratings	Clarity	The descriptions down here[referring to ratings] need to be very vague because it may be that the child is not going to anything and that is absolutely appropriate, say you'd have zero. It may be that the child needs five days a week childcare and is just not getting there and that may be a four; it's not about attendance it's about appropriateness that the child is getting. … Don't make it the absence of, make it the inappropriateness of [attendance]
Scale 13	Title	Clarity	Problems with appropriate engagement in non—parent care
Scale 14	Wording	Clarity	Change to: *problems with capacity to understand or knowledge to appreciate the infant's difficulties*
Scale 14	Wording	Clarity	Put *consider* instead of *include* the whole way through
Scales	Global	General	It's a lot better, I'm in the fact that we can even give specific micro-feedback speaks to the fact that the scale is a lot tighter… The one we tried 12 months ago was a lot more ambiguous then
Overarching question	Global	General	Does some places, doesn't in other places… A lot of work has been done, and we really do appreciate it [from previous version]

Group 4 - recommendations include to remove Scales 3,7, 11, 12,13

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
General	Education and training or instruction	Subdivide age range	Concern: The age range is so wide; looking at it with respect to a baby that five months old, it's really hard to score a baby that age according to some of these parameters because a lot of its written towards a baby that is much older with intentionality things
Index of scales	Scales	How to rate	At that age, would all non-accidental self injury be [Scale] 9?
Index of scales	Scales	Developmental stages	[Suggestion]: a couple of sentences under each one [each scale] for infants under 12 months, or under six months
Index of scales	Scales	Developmental stages	Examples specifically broken down into different ages – look at that category when you're looking at an infant at that age, rather than overall … Example Scale 5, language and communication difficulties, I was looking for something to express gaze aversion which is a predominant mode of communication for babies, and I don't think I saw one – it would be difficult then to express according to what's written there
Scale 5	Scales	How to rate	I wonder for infants in particular whether the language should be in a category of its own separate from other specific developmental delays… we see specific issues where the babies aren't verbalising; quite different from the other developmental delays, and for them all to be locked in together…
Index of scales	Scales	Structure	Eye gaze should be in Scale 9
Index of scales	Education and training or instruction	Subdivide age range	With babies it's much more difficult to know what's the cause of that, or does it fit in… It's much more difficult to go "that category's been dealt with" because if you've got a baby who's been screaming all night because of emotional dysregulation, you still want to categorise that as a sleeping difficulty, so a lot of the same things do actually come under lots of categories. Its less clear cut.
Index of scales	Scales	Structure	Issue: overt displays of distress and shutdown behaviours go together and I think for babies in particular they're quite different things that you're dealing with… should separate them
Index of scales	Education and training or instruction	Developmental stages	Also thinking about the young babies… because a lot of our patients in here are young babies, the mother cannot report on these things; she doesn't know that gaze avoidance for instance is abnormal and she can't report on that. So I can see it here in the session, but she can't report on what's been the problems in the last two weeks

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Index of scales	General	Usefulness	Also at that age, the baby's behaviour changes rapidly, so I suppose in that way the HoNOSI does pick up what I want, because between this patient's admission and the discharge [case mentioned] her relationship with her baby changed a lot, and so the infant's behaviours towards her did change that way as a rating of what's happening, it is worthwhile, so they're are pros and cons I guess.
Scale 1	Title	Clarity	Take out disruptive behaviour and just say "problems with disruptive behaviour"; it feels wrong to say that a baby is being disruptive, and I don't think we need to; an infant couldn't score a 4 on that scale, I imagine [agreement] [or]even a 3 [baby could not score]
Scale 1	Rating	Description	Hiccup and yawning could be misinterpreted differently – "yeah they yawned, there's a problem"
Scale 1	Rating	Description	[point raised by clinician to keep a six month old infant in mind as we go through the scales] "Turning away from all eye contact" [sounds made, indicating this is not appropriate]
Scale 2	Education and training or instruction	How to rate	I don't think *startled reflexes* is a sign of activity
General	Education and training or instruction	How to rate	[discussion not to rate according to cause – need to make clearer in preamble]
Scale 2	Scale	Description	Is underactivity written enough? - a child who is under-compliant, very good, I don't know whether that's fully in that - Not as much emphasis on the underactive, the over-compliant, overbright, good sort of child - to pick up dysregulation from that level
Scale 3	Education and training or instruction	How to rate	0-6 month old: at that age "distress"; might scratch. Have to rate as 9 [no problem/not applicable]. The issue is if a whole age group is rated as a 9, the scale is not going to work. The way it is, you can't do anything but give it a 9; or there's no problem, but there is a problem because the baby's distressed, but they're not self-harming.
Scale 3	Scale	Structure	… That for me feels like its been put in because its in the HoNOS and the HONOCSI and it really isn't applicable, and there are some 3 or 4 year olds perhaps, but that would fit in another scale anyway, and it sh*ts me, that sorts of thing, and I

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			think ethics committees would have a major issues if you were going to use that scale – the idea of raising that babies will self-harm …when she jumped of the height, she was looking for the ball, rather than injuries [laughter] …I really think it's a major issue
General	Global	Length - delete or integrate scales	[suggestion] – remove the whole scale- scale 3
General	Global	Length - delete or integrate scales	In fact, in terms of working with mothers and babies, that has been the sort of feedback we have discussed here, that the whole thing is so big, its daunting, and so if we were looking to simplify it to 12 scales, maybe, you know, so that's partly what's been on my mind. There's some that we could put into each other, or remove. …and if there are those older kids who are doing things, they'll be covered in other scales
General	Global	Length	That was my first thought when I first saw it; the size of it, and if practically people are going to use it, but its too big, it won't be used
General	Global	Length – feasibility, practicality	Yeah. That is what happens in here. We've got long experience of using scales and then stopping using them because they're too complex. Do you know Stephen Matthey's scale? From Sydney? Its an adaptation; there's a scale by Anton Guedeney, a French scale which is 8 items and 4 points, so its 32 altogether and that was too much, 8 times 4, was too much for people to fill in – it's the Alarm Distress baby scale; and then Stephen Matthey who is in Sydney adapted it, its used worldwide, and he simplified it, its called the modified ADBB – Alarm and Distress de bébé and his is 5 points – pneumonic FEVER – Facial Expression, Eye Gaze, Vocalisation, Activity Level and Relationship to Examiner and his is only 10 points altogether as against 32, because its 5 parameters and you can only get 0,1 or 2, so we actually do that as a screen for all of our infants here, so we discuss it at Ward Round and because its 10 point; that's the maximum that they can get, its easy, and there's only the 5 parameters, so its actually working quiet well in here. We had a scale we developed ourselves long ago, which was mainly based on the mother's maternal health rather than about the infant, but it was about the relationship, and even though it was developed here (the Louis Macro), we abandoned

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			it. The nursing staff were the ones who had to fill it in – that's it, it stopped. Even though we've got very good nursing staff and they're diligent, compliant, good people. That's a practical demonstration overall that FEVER's do-able, but its *minute* compared to this, its *minute!*
Scale 4	Scale	Description	Look at eating things they are not meant to eat; ingesting things other than food eg excrement or chalk
Scale 4	Scale	Description	And also overeating isn't there and that can be an issue too
Scale 5	Scale	Structure	Language and Communication in babies is different in other developmental delays and worthy of looking at in its own right…and other specific developmental delays [2 separate issues]
Scale 5	Education and training or instruction	Developmental stages	Age-appropriate vocalisations…you might need to have more examples of what they should be doing, because a lot of people won't know
Scale 6	Education and training or instruction	Developmental stages	Two infants in mind – 9 months and 9 weeks – in the 9 weeks baby - I don't think I could tell whether there was some language, in the time that I saw an infant
Scale 6	Global	Subdivide age range	Small infants are much more susceptible to tiredness, those sorts of things than an older child, so you'd expect less of an ability to say definitively "this is a problem" …particularly with one-off – where you have an hour
General	Global	Subdivide age range Developmental stages	See the Louis Marco's they were done with nursing staff over the whole week period and with a lot of 24/7 [much agreement heard]…exactly, and so when you're wanting people to do this and it will mostly be in an outpatient setting, I think its hard to assess with a baby …I agree …but its also very important
Scale 6	Global	How to rate	Do the colic-y babies get mentioned in here? …in the feeding one [Scale]…it includes…gastrointestinal symptoms …it doesn't just fit in one [Scale] …seems cleaner in the adult one…more clear-cut
Scale 7	Scale 1 and 7	How to rate	I think this is hard to tease out from Scale 1 for infants [probably until they are talking, or at least the under-one's], because you're seeing the same thing …could fit in hyperarousal – where does that fit?

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			…exaggerated startled responses, or floppiness – where does that go with emotional regulation, as well …its close, isn't it?
Scale 7	Education and training or instruction	How to rate	This is the difficult one
Scale 7 and 4	Scale	How to rate	Doubling up with feeding: *At more severe levels, the infant may be underweight and presents with failure to thrive due to discomfort and distress of feeding. Problems with sensory regulation may manifest as emotional dysregulation or emergent developmental delays.*
Scale 7	Global	Length - delete	This is a tricky one I think. …scrap it? Yeah! [much agreement]
Scale 7	Scale	How to rate	If it was a specific issue, I am thinking OT-wise, that could go under specific developmental delay; OT's look at sensory integration issues – over-reactive or under-reactive …the two OT's who work with us would be marvellous to review this tool
General	Education and training or instruction	Specialist Knowledge	If this tool is to be used by many, we don't want to make it something that can be used by someone with a specialist point of view – that's a problem I think …OT to advise on wording around sensory issues [for the tool in general]
Scale 7	Scale	description	Include sensory auditory processing issues
General	Global	description	It would be useful to go to a student perspective and at the opposite end, a specialist one …get the medical students to do it, and the clinicians
Scale 8	Rating	How to rate	…everyone gets a 4
Scale 8	Scale	Description	Snoring would be physical?
Scale 8	Scale	General	Scale is good
Scale 9	Description	Clarity	[scenario described where a child was removed by protective services]…and everybody said "Isn't he good" but his reaction was *extreme*…and it was severe…overly controlled…internalisings…child is at enormous risk – [scale] does not describe the child
Scale 9	Description	Clarity	This scale is a mish-mash
Scale 9	Scale	How to rate	Acting out or shutting down should be two separate scales for this group rather than covered in both
Scale 9	Scale	How to rate	Shut down to me means a different thing, than a good child doing the appropriate behaviour …I don't feel this tool captures this in any way

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 10	Scale	General	Like it [Scale] ..that one's all right
Scale 10	Education and training or instruction	Specialist knowledge	Level of knowledge of the clinician [describing scenario where a child gives the clinician a hug at the beginning or end of the session and an inexperienced clinician may think that they are being social, but it could be a problem because it's a relief; not being social – need to recognise that; disorganised attachment, who do the indiscriminate social interaction]
Scale 11	Scale	How to rate	Another one that's harder with the smaller babies
Scale 11	Scales	Structure	*Problems with separating from parents when the infant is attending structured socialisation settings* …I find that more of an emotional issue, its not about environmental exploration
General	Global	delete	Wonder whether it could be coalesced into something else? …I wonder if it can be scrapped? [Agreement heard]
General	Global	delete	Because of all the excluding suggestions [Scales, 8,12,13], I think this is covered in those
General	Global	delete	I think this is one of those that's filtered down from the HoNOS, then the HoNOSCA – I think that's one that isn't really relevant …if everything else is OK, this is going to be OK
General	Global	delete	Scale 10 – social reciprocity – if the parent's doing that well, in terms of exploration and secure base, which is social reciprocity, then [Scale] 11's going to flow on, you could almost say it could be the behavioural manifestation of what's happening
General	Global	delete	…and if its not [going to be OK], then it could be covered in other scales
Scale 12	General	Scale	Fine
Scale 12	Rating	How to rate	Score 4 [patient scenario described] – mothers' view of the child rather than her behaviour towards him – this may not capture the child [suggested change to:] "*Likely* to have severe impact" [where there isn't a severe impact] – *Would be* likely, because we can see that - should be coded as a problem – risk, potential impact – if we change the wording, it could improve it
Scale 13	Global	Structure	This is another one that has been brought down from the HoNOS and the HoNOSCA – its about whether the child gets to go to child care and things like that – sometimes they're made to go even if it is a real problem being there, because they get taken there, even if attending is an issue – this is taking the school attendance idea - a 6 month old can't refuse to go to child care

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 13	Global	Delete	I just actually think that should be dropped
General	Global	Delete	Down to 11 [Scales]
General	Global	Delete	Its of marginal value to this …I think scrap it! …it just adds to the burden of doing the scale if you have to do it every time and its not going to be relevant to the vast majority …and it doesn't matter if they don't go to things because they don't *have* to go to things, its not like a school aged child not going to school …kids don't have to go to child care …it could be included in Scale 15 *a lack of understanding about services or managing the infants' difficulties* – it could well be included - you could *easily* just say "Parent has difficulty getting the kid to child care when the child objects, not seeing the value of child care for the child, child needs limit setting" – you could *easily* incorporate it into [Scale] 15.
General	Global	Delete	Its just feels like it was put in there to fit in with the others
Scale 14	Wording	Clarity	Lack of useful information is a bit of a furphy [suggested its not needed] -how would you know that? Not having been given the information is different from not having the capacity to take it on and support the infant appropriately
Scale 14	Scale	General	Scale is all right
Scale 14 and 15	Scales	Delete/ consolidate	Problems with lack of useful information – could come into 14 – collapse them together
Scale 14	Scales	Structure	14 and 15 are both there, but its not *adding* anything to the assessment, because one flows from the other
Scale 15	Scales	General	Another one that has been brought down from the other scales and its from a disability model in the HoNOS. What might be more relevant is that parents aren't accessing the supports and services that would be helpful for their infants, or its not available
Overarching Question	Global	General	I think is we looked at it as being a before and after measure and we're looking if there's improvements in their mental health, possibly not, given what we're talking about internalising and externalising stuff
Overarching Question	Global	Global	I think it would capture that and I was thinking to put that to the test, so maybe when we have new patients [discussion with a staff member] you could do a before and after, because we haven't done it accurately enough to give that, for here.

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Overarching Question	Global	Global	Do you think it would work in a CAMS setting? …if its not too targeted and narrow, that there is enough – the skill level is very different to make an assessment with young ones in CAMS. At the moment it would be difficult for a CAMS clinician …but in a way, a scale like this could be helpful, because it is more explicit, so it can help to define out some of the abnormal behaviours
Overarching Question	Global	Global	Two overall comments: it's a very wide age range: 0-9 or 12 months is different from 1-4 [years] or even 0-2. If you make the babies too wide, you're not going to catch the baby stuff, that's why I was suggesting 12 months
Overarching Question	Global	Global	I think its going to be really hard, even here, and I say even here, because we've got really expert staff and we're relatively well resourced, its hard to do a scale of this complexity and you just get buy-in resistance
Overarching Question	Global	Global	It's a good tool for general CAMS to use to point them in the right direction of thought, what to capture, what to observe, what to think about, so I do think its useful …would hone down the thinking – the focus needs to be in not just developmental difficulties in the child, but something parental

Group 5

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Key principles	Key Principles	General	Clear if you've used HONOSCA – looks like a downwards revisions of that …Fairly self-explanatory
Index of scales	Scale 1 and 9	Structure	This is the one that I had the most difficulty with, working out the crossover between Scales 1,2 and 9
Scale 1	Education and Training or instruction	How to rate	This is the one that I had the most difficulty with in working out the crossover between this one and number 7. I thought there was some crossover between Scale 1,2 and Scale 9 Scale 9 talks about exaggerated positive and negative emotional response and some of the examples given in Scale 1 – *demanding, irritability* - I think it might be difficult to know which ones, where to score some of those things.
General	Education and Training or instruction	Developmental Stages	I think what we need is for people who *don't* do a lot of work with infants and would be expected to – there needs to be some kind of reference of what's to be expected for normal development – that's just my experience. At the Children's Hospital, a number of clinicians will see infants and they're not grounded in very early infant development, so they might say "What is an exaggerated positive or negative emotional response?" How do you define exaggerated? What are the parameters around that for someone 0-3 months, 3-6, 6-9 – there's a lot that goes on in terms of very early development. Then when you've got the more disruptive behaviours with the toddlers …what does it look like in a 3 ½ year-old versus a 2 year old …what's the spectrum? I do think that that's a frame of reference that's required. Unless you work with young children all the time, that's not what people are experienced in and know a lot about …it does say that working at this level of CAMHS, you would be expected to have this level of knowledge, but I think that's probably unrealistic …particularly for people working in the country – there are some very remote CAMHS, who don't have a lot of contact with other professionals. …I think you're right as the rate of change is so enormous – 3 months versus 9 months or one year, so you've got to know …says something about the *capacity to manage intense feelings* [Scale 1] *of hunger…*well a young baby can't manage those intense feelings, and

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			that's normal, so I agree that's difficult; there might even need to be some caveats
General	Education and Training or instruction	Developmental Stages	Bold: "Age and developmentally appropriate capacity", so if they don't know, they can go and find out." …needs a developmental glossary …just a brief account of what's to be expected at each particular age, I suppose
Scale 1 and 9	Scales	Structure	Scale 1 is about disruptive behaviour and Scale 9 is more about introverted sort of behaviours; depression, anxiety, phobia; I suppose that doesn't really take into account there might be underlying feelings of one set or another which one could regard as internal, but of course it manifests externally, but I don't know, I think that's making it far too complex and probably to just see those two scales as reflecting different aspects of behaviour problems – seems to me they do
Scale 2	Title	Clarity	*Problems with joint attention* and then later on Scale 10 is *Problems with Social Reciprocity*, which in part of social reciprocity is having joint attention and I wondered how useful the word *joint* was? …confusing…made me think of developmental issues …confusing …if you took out the word joint, it would be fine…*sustained attention and care-giving relationships,* or something
Scale 2	Description	Clarity	Exclusion – need restlessness, jerkiness as they might be having medical difficulties …that applies across the board to all the scales…even in very young prems, you have no idea what the impact of that is really
Scale 2	Description	Clarity	Physical illness or disability – few overlaps there … Either make it an exclusion, or you need to make it more explicit, that you include
Scales	Global	Clarity	Clarify that you're rating observation, not cause
Scale 2	Description	Clarity	Is it model of attention? For example, Mirsky's [model] is famous and endorsed and it focuses on attention, selective attention, divided attention – if you want to go down to problems of attention, how neuropsychological do you go? Add descriptors. …could be more of a red herring…there is more than joint and sustained attention in those models – for people who are not familiar with those models, we were talking about autism, social reciprocity … *Joint* is not a term I am familiar with

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scales	Global	Specialist knowledge	I doubt this will be used for children under two, in the present situation… The point is you've got to make it understandable by someone who doesn't very often use it
Scale 3	Description	Clarity	You could perhaps bypass the *attention* issue by using descriptors such as *concentration* and *focus* instead
Scale 3	Description	Clarity	Clear. I thought its strength was that it's really clear to know what to include. Eg. don't worry about intentionality – I think that's useful. …its well done
Scale 4	Description	Clarity	What's *All glossary examples should be developmentally referenced?* … That's my question too. … Global comment? Specific to feeding?
Scale 4	Description	Clarity	While I think it's important to include feeding problems related to physiological problems, sometimes there's physiological problems that can be changed if there is an underlying psychological issue as well compounding that, but sometimes it doesn't matter what you do, there's not much change - if we're looking at this as a measure to see how change is unfolding over time, some things won't ever change. … Exclude children where there is an organic cause due to feeding difficulties that are unlikely to change due to the congenital nature of the problem.
Scale 4	Rating	Clarity	[Rating 3] *some risk of nutritional problems* and [Rating 4] … *A high probability of severe nutritional problems* do these need to be here? … Even a fussy eater may have nutritional problems … Does it include the obese infants? Does not seem to refer to obesity at all, which surprises me, given the current emphasis. They lay down fat cells and never lose them, so prognostically it's a bad thing. … I would have thought it should be included … It's got the mental relationship between hunger, feeding and saiety … Include under-nutrition *and* over-nutrition
Scale 4	Title	Clarity	*Problems with feeding* and eating *behaviour* to signify it does not just refer to babies
Scale 5	Scale	General	Broad. What to consider here.
Scale 5	Wording	Clarity	*Pragmatics* – what does that mean? … Quite sophisticated some of the language that's used. Use less jargon, give an example.

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 5	Description	Clarity	Why have you got the language or communication difficulties outlined but not some examples of developmental delays?
Scale 5	Title	Clarity	ordering to be different as well [title]. For example, *problems with specific developmental delay in language or communication* because that implies language, communication and other specific developmental delays
Scale 5	Scales	Where to rate	*Cognitive, motor* - I read that and thought: do they have a *motor* developmental delay? … Which I think scale 6 comes into play as well … So we're just rating language and communication in this one?
Scale 5	Description	Clarity	For example a child with Cerebral Palsy do you include them or exclude them, because it says exclude those with a disability, but in fact you include them as they've got motor problems. … Needs to be more detail about what specific developmental delays are included with examples of motor or coordination difficulties; examples are all language-based, so it makes it feel like that's all you're interested in …its confusing with Scale 6 about physical illness or disability …clarify motor or other specific developmental delays include…
Scale 6	Description	Clarity	Similar to our comments on the previous one – motor developmental difficulties – where to put those
Scale 6	Wording	Clarity	Change *drug and alcohol use* to *drug and alcohol exposure*
Scale 6	Description	Clarity	Those physical illnesses such as epilepsy – are they captured here?
Scale 6	Wording	Clarity	Language needs tidying up … *Movement disorder…* Then you've got *Side-effects from medication* which is completely different area yet again, *then congenital…*how its presented; – that's a lot …examples of movement disorders … *Physical health problems as a result of psychological issues*…I'm racking my brain – could be more clearly presented - make more explicit - physical health problems with no identified organic cause …its messy …random *physical complications of psychological disorders such as severe weight loss* - why isn't that in feeding behaviours? If it's in an infant, it's not like they have anorexia

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			… Severe weight loss [should go in scale 4] nutritional problems … Examples need to be a lot clearer and presented in a logical fashion
Scale 6	Rating	General	Descriptions of the ratings are quite good. I feel I could rate them if I was clear about what I had to include.
Scale 6	Description	Clarity	The opposite has happened with *sensory processing* it's taken a lot of work to explicate to what that means
Scale 7	Description	Clarity	Good that there is a lot of information included for people who are not familiar with this field, it needs to be spelt out
Scale 7	Wording	Clarity	*Problems associated with the regulation and integration of sensory processing* [delete:] predictable
Scale 7	Description	Clarity	Did not like *for example, the question is not whether their hearing is OK, but rather, is their processing of auditory information peculiar or problematic?* Use visual, tactile, sensory – need examples; how its experienced. Examples of good *manifestations of…* That kind of everyday language would be more useful.
Scale 7	Wording	Language	Do not use *OK* - sounds too familiar. Rest is technical, this is too familiar.
Global	Scales	Consistency	It seems like certain specialists have contributed to the development of each of the scales… You get the sense there are slightly different styles, a little bit more detail in one and not the other… Ideally would have the same flow of language and similar level of detail for each scale
Scale 7	Rating	Clarity	Descriptions of each rating were good
Scale 8	Education and training or instruction	Developmental stages	Reminder of what is developmentally appropriate at each age – because that does really change eg. baby's not sleeping at 10 weeks - that is totally fine
Scale 8	Wording	Description	Change *disturbance* - I see "problem"; threw me off
Scale 8	Education and training or instruction	Developmental stages	Repeat in ratings - in keeping with age expectations
Scale 8	Description	Clarity	Co-sleeping should be included but not put in a way that it pathologises it - open to interpretation … Could be cultural – may not matter – still a problem … Put in HoNOSCA and all scales
Scale 8	Rating	Clarity	[Rating 4] include *age "*and cultural" *norms*

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 8	Description	Clarity	Included night terrors but that is very normal… Maybe none of them... Parasomnias… need to be included [ie. Exclude them, esp night terrors … Capacity to settle independently should be included
Scale 9 versus 1	Global	Where to rate	Transposing emotion and related symptoms… You could almost plagiarise the headings in ICD 10, to put in this scale
Scale 9 versus 1	Global	Where to rate	Back-arching - scale 1 – overlap
Scale 9	Scales	Where to rate	*Settling, demanding, wining, irritability, excessive crying* – that whole paragraph would probably be better in scale 9 because they're not disruptive behaviours; they're signs of a child distressed, whereas biting, hitting, being defiant – they're behavioural
Scale 9	Scales	Where to rate	Haven't got non-organic symptoms here. Eg. if a child is still soiling inappropriately, not due to physical or mental impairment, could be a sign of rage - that would be in the disruptive behaviour one
Scale 9	Description	Clarity	Confusing – exclusions – *physical sequelae of psychological disorders or medication* you are going to have physical sequelae in young infants; that's how they show distress … Medication maybe, but the other bit does not make any sense … Trying to think what that might be
Scale 10	Title	Clarity	Add: *and interpersonal sensitivity*
Scale 10 and 2	Description	Clarity	Joint action better placed here rather than scale 2?
Scale 10	Education and training or instruction	Specialist knowledge	What is normal eye gaze – infants do turn away to regulate, come back after regulating include the word *sustained* avoidant gaze
Scale 10	Wording	Clarity	*includes problems with* "parents" *emotional attunement* "and their responses to infant's needs"
Scale 10	Description	Clarity	Including more paragraph breaks – helpful when you're reading things [referring to the last paragraph of inclusion box on scale 10]
Scale 11	Scale	General	Routine … Fine
Scale 11	Description	clarity	would not say *it is critical to rate the problem and good practice to not assume* I would just say: it is vital to pay attention to the fact that infants… Something like: Infant's difficulties may be related to the parent's capacity …later confused… *Do not include lack of opportunities for exercising intact abilities and skills*

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			as might occur in an over-restrictive family rated at scale 12 - how that overlaps with what we've just been talking about – helicopter – need the caregiver to scaffold it for you in order to be able to do it
Scale 11 and Scale 9	Scale	Where to rate	Self-care does not belong here. Should it be independent from environmental exploration – if important, then paragraph break, or separate Scale for environmental exploration [Scale 9] … Two quite separate areas
Scale 12	Scale	General	These issues are done better than the HoNOSCA eg irritability. Scapegoating – nowhere neat to put that in the HoNOSCA
Scale 13	Scales	Where to rate	Some of that stuff that included exploration may be doesn't need to, if it's done here
Scale 13	Scales	Clarity	Big section may not be warranted. If not going to a childcare centre … its about can they cope in a social situation apart from the one of their primary care giver … Focus on educational care settings feels like its done to stick with item 13 on the HoNOSCA, rather than what's developmentally appropriate
Scale 13	Scales	Delete or split scales	Socialisation is elsewhere and Anxiety is elsewhere so you actually don't need it …so ditch it …and you could split one of the other two … Not have that one but an elaborate on one of the other ones – developmentally does not fit with age group [school and backpacks mentioned here]
Scale 14	Scales	Clarity	I never understand the difference between 14, 15
Scale 14	Education and training or instruction	Stages of development	It would be interesting to rate what parents understanding is of normal development in the child - if parents don't understand what is typical and they need developmental guidance, that is a problem – a subset of understanding their child's difficulties
Scale 14	Wording	Clarity	*Include problems with capacity or knowledge to understand the infant's difficulties* add "and needs"
Scale 15	Description	Clarity	*Lack of understanding of the most appropriate intervention* – should be explicated that it's not parents' lack of understanding, its systems lack of understanding
Scale 15	Rating	Clarity	Give more examples of *moderately severe* and *very severe*, like in scale 14 to explain how they are different
Scales	Scales	Clarity	[scales 14 and 15] I would be interested to know why these were included originally; how they got to be there

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Overarching question	Global	General	Partially to mostly …mostly is a good word
Overarching question	Global	General	With improvement, it would be much more useful than the HoNOSCA. Once you've finished with this one, can you review the HoNOSCA for Primary School ages children, 5-12 year olds?

Group 6 – Specialist CAMS – Infant MHS, not PIMS. Specialised Infant Mental Health Team (CYMHS)

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Key principles	Key principles	General	Clear
Scale 1	Description	Clarity	*Under-controlled emotional regulation* is not in the description at all; it's all about aggression, which is an issue when we're looking at the little babies who generally aren't aggressive
Scale 1 and 9	Scales	Structure	Include under controlled emotional regulation …withdrawal comes into scale 9 …I think the context is also that it needs to match up with the HoNOSCA…but for us the struggle is how we bring the infant into this…our feeling is that it doesn't work so well
Scale 1	Scales	Structure	The little ones who aren't settled – we're not getting that into the picture
Scale 1	Scales	Structure	It's feeding, sleeping and socialisation that can be impacted by being under-regulated or externalising as an infant
Scale 1	Rating	How to rate	All about aggression/disruptive externalising
Scale 1	Rating	Developmental stages	[Rating 4:] *Disruptive or aggressive in almost all activities. At least one serious physical attack on others or animals, or serious destruction of property* How do you apply that to a six-month old or a four-month-old? Its not working.
Scale 1 and 9	Scales	Structure	Include frozen …some of this you will pick up in [scale 9] [discussion re the confusion of over-controlled versus under-controlled and which scale] … Frozen watchfulness is in [scale] 9 …add: do not include – see scale 9 …over-controlled comes up more in [scale] 9 …make it éxternalising" versus "internalising" …we changed that but no-one knew what it meant, so we change it to over- and under- controlled
Scale 1	Rating	Clarity	Age range is not reflected in the ratings …what this means in an infant …generic CAMHS are more likely to underrate the infants because they're not displaying aggression, so infants are going to come out underscored
Scale 1	Education and training or instruction	Specialised knowledge	The description is not going to point people in the direction enough, if they're not a specialised service… The reality is, across Australia, most people would be generalists [referring to infants]
Scale 2	Scale	General	Easier than Scale 1 – statements are more appropriate across age group
Scale 2	Scale	General	Happy with [Scale] 2 …its about scoring the infant, not the relationship

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 3	Description	Clarity	Category that is very difficult with the little ones that we would see… It's hard to keep the real tiny ones in mind when you read this
Scale 3	Description	Clarity	Self-soothing behaviour – need examples for infant
Global	Rating	Clarity	Generally I think the descriptions are adequate… Is to slowly move down to each category [ratings], they don't necessarily even reflect what's in the description …[Rating 4] *Overt inhibition of response to pain/discomfort and lack of self protection and self soothing* does not match the descriptor at the top.
Scale 3	Description	Clarity	*Making dolls hits their heads during playing* I'm wondering if that's more kind of fantasy, rather than self injury? You could say rough play, not self-injury [discussion on what self injury means in infants] … I don't know if this actually applies to babies, the way it sits here – when you've got the feeding one …12 months and under, struggled to see how it would fit [for under 12 months]…Excessive sucking, or skin scratching or picking [might fit], or throwing themselves backwards…would not happen often …0-6 month old – we may not see it at all …include for 12 months and under – *observations might include…*
Global	Description	Clarity	…for infants it needs to consistently be set out in each one [scale]
Scale 4 and 9	Education and Training or instruction	Where to rate	If eating to regulate – is that Scale 9?
Scale 4	Description	Clarity	*little recognition of the relationship between hunger, feeding and satiety* – may be a biological predisposition to over-eating, might be nothing to do with the relationship
Scale 4	Description	Clarity	Need a sentence to cover under- as well as over-feeding
Scale 5	Rating	Wording	*Mildly severe* – how can you have those together? Suggestion: mildly significant
Scale 5	Description	Clarity	Add: *fine and gross* motor…[first line]
Scale 5	Description	Clarity	Confusing…*don't include infants with global learning disability*…but then up the top its say to put that in, irrespective of the cause or solution
Scale 5	Description	Clarity	Less than 12 months – do not know if there is a global disability
Scale 5	Description	Clarity	Bigger emphasis on language but it seems like you're looking at all areas of development - there's

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			a lot more examples of speech and communication rather than any other kind of development
Scale 5	Description	Clarity	Include social development …scale 10 is on social reciprocity
Scale 5	Title	Clarity	Change title [no suggestion]
Scale 5	Description	Clarity	Include examples of slow motor development
Scale 5	Description	Clarity	Is this only about delay, or also include precocious development
Scale 5	Description	Clarity	Include problems with slow processing – important to observe that
Scale 5	Description	Clarity	Include limited use of words to express themselves; talk about expressive language …do you think the people using this would be all be familiar enough that you could use words like *expressive* and *receptive*? …include problems with understanding producing (communicating) and expressing language
Scale 5	Wording	Clarity	Word *pragmatic* needs to be explained
Scale 6	Description	Clarity	*Physical effects from drug and alcohol use* – needs to be more clear …what about people who give their kids drugs? [Phenergan/Panadol; any antihistamine with sedative properties] …or all of our parents who dipped the dummy in the brandy or something! …would be dopey when you visit them as a side effect of medication …could mean from the parent or the child being given them
Scale 6	Ratings	Clarity	[Rating 2]: *the infant has some mild symptoms of physical illness or disability, that occasionally prevent engagement* add *prevent or challenge engagement*
Scale 6	Ratings	Clarity	[Rating 4]: *not possible* – change to "significantly or seriously impacted"
Scale 6	Description	Clarity	Change *physical effects from drug/alcohol use* to distinguish between ingestion versus being given it
Scale 7	Education and training or instruction	How to rate	The difficult thing with diagnosing a sensory problem is that we don't know whether it's a biological/emotional problem or due to the child being emotionally dysregulated …not aetiological – its what's happening with the infant …not talking about diagnosing, but what factor it has on the child's health … May indicate emotional/biological difficulties …biological rated at Scale 5 rather than here? … Clarify that it may be associated with underlying difficulty processing their emotions

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			…could be indigenous to the child or their environment; rate both here – as the second sentence in paragraph 1 of description
Scale 8	Description	Clarity	Weighted very much on parental ability to cope with it – does not relate necessarily to the sleep problem. You may have a very relaxed parent who couldn't care less, or you might have an uptight parent who can't cope with one waking a night. … Use some measure other than whether it interferes with the family
Scale 8	Education and training or instruction	Developmental stages	Add a sentence regarding parental expectation versus expected parameters of this age and developmental stage
Scale 8	Rating	Clarity	Does it need to be that its a rating of 4 before it impacts on the child's development?
Scale 8	Rating	Clarity	*The sleeping pattern is a cause for great distress in the parents and* "interferes with the child's functioning" or *something*.
Scale 9	Description	Clarity	Add *trauma* after *fear* in first paragraph
Scale 9	Description	Clarity	Add *trauma* after *fear* in first paragraph
Scale 9	Description	Clarity	[NB: – detailed, specific discussion to get these words right:] Change to: *"Include age or developmentally inappropriate lack of appropriate comfort-seeking or safety-seeking, lack of wariness or avoidance of adults/others, and disinhibited behaviour"*
Scale 9	Description	Clarity	Include age-inappropriate sexualised behaviours with self or others
Scale 9	Description	Clarity	After *curiosity* "decrease in exploration and play"
Scale 9	Rating	Clarity	Jumps seem excessive – 2 to 3, then 3 to 4 don't seem even enough
Global	Ratings	Clarity	Wording: Only bring in *severity* at rating 4 – make more consistent
Scales 9 and 10	Scales	Where to rate	It's got *indiscriminate and overfamiliar social interactions* in this scale as well – do we want it in this one and the one before? Now that we've just spent all that time adding it in to the one before? I think there is a difference between disinhibited to a social interaction… Disinhibited can be an emotional type response. Is indiscriminate different to disinhibited? … At risk of rating it twice
Scale 10	Description	Clarity	*Eye contact* also use social referencing, to check what someone is thinking or feeling Change to *eye contact or to socially reference others*

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
Scale 5 and 10	Scales	Structure	Pragmatics here, rather than [Scale] 5 Understanding, expressing and functionally using language
Global	Scales	Ratings	I prefer this type of rating where you obviously have to refer back to the explanation, rather than the explanation being embedded in the rating; I think that's easier …would make things easier …that's how the HoNOSCA's done
Scale 11	Title	Clarity	Add: *in an infant*
Scale 11	Description	Clarity	Difficult to assess in an infant. Eg 2-month old – I just rated as a zero
Scale 11	Description	Clarity	Add: Environmental expression in an infant may be expressed as visual curiosity, body movements to show excitement
Scale 11	Description	Clarity	Add: restrictions or conditions of environment in that second paragraph eg. kids who have been in detention
Scale 12	Description	Clarity	First paragraph – highlight also siblings
Scale 12	Ratings	Clarity	Ratings – OK
Scale 13	Description	Clarity	Didn't really apply in an infant still at home. Add: Do not rate very young infants – don't want to give them a low mark because they're very young and do not go to play group …so that would be a 9
Scale 13	Description	Clarity	…include not just opportunities, but their ability to use opportunities
Scale 13	Rating	Clarity	[Rating] Opportunity versus reluctance - *refusal* due to anxiety?
Scale 13	Description	Clarity	Change to *Include limited or minimal opportunities* and capacity *to attend* and participate in *socialisation activities*
Scale 13	Title	Clarity	Problems with *participation* and attendance
Scale 13	Rating	Clarity	Make rating consistent: minor, mild, moderate
Scale 14	Scale	General	Great
Scale 14	Rating	Clarity	Standardise the rating system
Scale 15	Scale	General	Outstanding – no need to be modified
Overarching question	Global	General	Not without changes …its getting closer …of those changes are made, then yes …if the over-two's it probably does it now, but for the little babies it needs modification …the under 12 months …We're really looking forward to using it …Overall, we would be really pleased to have something like this that we could use …it didn't actually take a massive amount of time; I think once you're familiar with it

Topic/Scale	Code	Theme	Focus Group Responses and Quotes
			...if we can make it work across the age range that we work with, it will be good.

Appendix 3: Transcription of Issues from Focus Group Data Organised by Theme

This section shows the raw focus group data transcriptions and quotes which have been organised by theme.

Appendix 3.1 - Theme: Reference to Developmental Stages

Group 1, page 11 (of original transcriptions)

Scale 2	Education and Training or Instructions	Reference to developmental stages	And this is with a talk about it going to be in age development, because I've actually seen earlier versions of this document and one of the issues will be what a 3 month-er can do, is going to be very different to what a 47 month-er can do, so therefore the person doing it needs to understand the difference between the two and I'm not clear that will happen
Scale 2	Education and Training or instruction	Reference to developmental stages	I think something needs to go with it. There needs to be some training, and it needs to be some reference to the different stages
Scale 2	Education and Training or Instruction	Reference to developmental stages	So maybe some recommendations about who is qualified, or who is benchmarked, whatever term they want to use, to know the difference between using the tool on the 47-month old and using the tool on a three-month-old
Scale 2	Education and Training	Reference to developmental stages	That's probably not that dissimilar to HoNOSCA in a lot of ways because you asked to rate 4-17, and if you don't have a sense of what those developmental stages should look like…

Group 3, page 2

Global	Structure	Specialist knowledge Developmental stages	… I agree. I think as we move into the instrument, the severity ratings become less wordy and I think it's kind of clearer that the clinician makes a judgement about mild, moderate or severe, rather than trying to find examples, because people get caught up on the examples, I think. Clinicians do when they're rating it, they think "oh no, I don't have that particular thing", but they maybe missed some of the different… What I like about it, one of the improvements from the previous versions we've tried, was that this a bit more in younger infants and preschoolers,… Divide… Which I think is helpful for clinicians, particularly if they're not as confident with younger infants, to have some examples of how things might present in an infant.

			I think the examples of there, it might just be about structuring them in a consistent way, for younger infants and older. [Generalised comments to put examples in descriptor, not rating]

Group 3, page 6

Scale 13	Education and training or instruction	How to rate	This is an interesting one for infants. This is the equivalence of school attendance/work [referring to scale being derived from other scales] … In infants, is that about separation with other caregivers? What symptoms are pathological in a three-month-old difficulty engaging in child care? I think that's the challenge; what the symptoms are.

Group 4, page 1

Index of scales	Scales	Developmental stages	[Suggestion]: a couple of sentences under each one [each scale] for infants under 12 months, or under six months

Group 4, page 1

Index of scales	Scales	Developmental stages	Examples specifically broken down into different ages – look at that category when you're looking at an infant at that age, rather than overall … Example Scale 5, language and communication difficulties, I was looking for something to express gaze aversion which is a predominant mode of communication for babies, and I don't think I saw one – it would be difficult then to express according to what's written there

Group 4, page 1

Index of scales	Education and training or instruction	Subdivide age range	With babies it's much more difficult to know what's the cause of that, or does it fit in… It's much more difficult to go "that category's been dealt with" because if you've got a baby who's been screaming all night because of emotional dysregulation, you still want to categorise that as a sleeping difficulty, so a lot of the same things do actually come under lots of categories. Its less clear cut.

Group 4, page 1

Index of scales	Education and training or instruction	Developmental stages	Also thinking about the young babies… because a lot of our patients in here are young babies, the mother cannot report on these things; she doesn't know that gaze avoidance for instance is abnormal and she can't report on that. So I can see it here in the session, but she can't report on what's been the problems in the last two weeks

Group 5, page 5

Scale 5	Education and training or instruction	Developmental stages	Age-appropriate vocalisations…you might need to have more examples of what they should be doing, because a lot of people won't know

Group 5, page 5

Scale 6	Education and training or instruction	Developmental stages	Two infants in mind – 9 months and 9 weeks – in the 9 weeks baby - I don't think I could tell whether there was some language, in the time that I saw an infant

Group 4, page 5

Scale 11	Scale	How to rate	Another one that's harder with the smaller babies

Group 5, page 1

General	Education and Training or instruction	Developmental Stages	I think what we need is for people who *don't* do a lot of work with infants and would be expected to – there needs to be some kind of reference of what's to be expected for normal development – that's just my experience. At the Children's Hospital, a number of clinicians will see infants and they're not grounded in very early infant development, so they might say "What is an exaggerated positive or negative emotional response?" How do you define exaggerated? What are the parameters around that for someone 0-3 months, 3-6, 6-9 – there's a lot that goes on in terms of very early development. Then when you've got the more disruptive behaviours with the toddlers …what does it look like in a 3 ½ year-old versus a 2 year old …what's the spectrum? I do think that that's a frame of reference that's required. Unless you work with young children all the time, that's not what people are experienced in and know a lot about

			…it does say that working at this level of CAMHS, you would be expected to have this level of knowledge, but I think that's probably unrealistic …particularly for people working in the country – there are some very remote CAMHS, who don't have a lot of contact with other professionals. …I think you're right as the rate of change is so enormous – 3 months versus 9 months or one year, so you've got to know …says something about the *capacity to manage intense feelings* [Scale 1] *of hunger…*well a young baby can't manage those intense feelings, and that's normal, so I agree that's difficult; there might even need to be some caveats

Group 5, page 2

General	Education and Training or instruction	Developmental Stages	Bold: "Age and developmentally appropriate capacity", so if they don't know, they can go and find out." …needs a developmental glossary …just a brief account of what's to be expected at each particular age, I suppose

Group 5, page 5

Scale 8	Education and training or instruction	Developmental stages	Reminder of what is developmentally appropriate at each age – because that does really change eg. baby's not sleeping at 10 weeks - that is totally fine

Group 5, page 5

Scale 8	Education and training or instruction	Developmental stages	Repeat in ratings - in keeping with age expectations

Group 5, page 7

Scale 14	Education and training or instruction	Stages of development	It would be interesting to rate what parents understanding is of normal development in the child - if parents don't understand what is typical and they need developmental guidance, that is a problem – a subset of understanding their child's difficulties

Group 6, page 1

Scale 1	Rating	Developmental stages	[Rating 4:] *Disruptive or aggressive in almost all activities. At least one serious physical attack on others or animals, or serious destruction of property* How do you apply that to a six-month old or a four-month-old? Its not working.

Group 6, page 3

Global	Description	Clarity	…for infants it needs to consistently be set out in each one [scale]

Group 6, page 4

Scale 8	Education and training or instruction	Developmental stages	Add a sentence regarding parental expectation versus expected parameters of this age and developmental stage

Appendix 3.2 – Theme: Overlaps and Crossovers in Scales

Group 1, page 7

Scale 1and Scale 10	Description	Clarity	On that paragraph where we're going more into specific examples, *Include difficultly*, I can see I think there appropriate specific example *undue irritability, excessive crying, frequently arching back*, maybe not so much in the sense of now you're getting really specific, whereas with excessive crying, undue irritability, people can imagine what it looks like, but anyway, going down to *inhibition of expected interactive response*, I found that one really vague and naturally linking more to other scales like the *social reciprocity* in scale 10, I am not sure what is meant by *inhibition of expected interactive response* so maybe they're trying to talk about not that the child doesn't sooth when we're expecting it to sooth, that's the *expected response*, but *interactive response*, is more we'd be expecting generally when we talk about interactive response, at least for me would be like the *social reciprocity*, more like the shared

			eye contact for infants, yes for their emotion regulation to settle, but they do that in a social way (other respondent:) what they're presumably meaning is that when an infant's attempted to be settled, and they're really not engaging with that, I mean that's kind of what its looking at (other respondent:) I thought something different. I thought its about the situation where the infant's actually inhibiting their response to an interaction with the caregiver rather than just naturally responding (other respondent:) then that would be in *social* [agreement heard re interpretation of meaning] (other respondent:) what she said! [ie indicating agreement] (other respondent:) I think that makes sense, but not in the context of this scale…for what they might have meant in this scale (other respondent:)that would almost be *over*controlled emotional regulation and that's why it doesn't seem to fit, but maybe that's not what they mean …I mean it doesn't actually ask you to rate that under emotion (other respondent:)but it does talk about *psychological indicators of stress (hiccups, yawns, non-injurious scratching)* (other respondent:) yeah. I really like that part

Group 2, page 3

Scale 2	Scales	Structure	Why put activity and attention together? You may have problems in one or the other domain and clinically if you have both, it's much more serious. I wonder if we should have two scales for those? I'm thinking about trauma and developmental disorders. I am thinking of kid who's got a very high activity level and very poor attention as opposed to someone who's got an active temperament but can sustain attention.

Group 2, page 5

Scale 1 and 9	Title	Clarity	Internalising – withdrawal, listlessness – reflect in the title, as opposed Scale 1. [Suggest:] internalising

Group 2, page 6

Scales 11,12,13	Scales	Where to rate	In general, I think there is a lot of cross-pollination over the next three items [Scales]; how environmental exploration interplays with family life, interplays with opportunities for socialisation, and to some extent I think it will be sort of artificial how we pick those things

Group 3, page 1

Scale 1	Scale Structure	Structure Title	We thought that this is something that perhaps should be separated, the disruptive behaviour from dysregulation… Certainly someone could be dysregulated, and nothing to do with disruptive … Yeah, and that would come up in other scales

Group 3, page 1

Scale 1 and 9	Scale	How to rate	Dysregulation should probably come under Scale nine… So just extract the dysregulation symptoms... and put them into the examples for the emotional. The dysregulation is described in the descriptor in a non-aggressive way, and then also in an aggressive way

Group 3, page 1

Scale 1	Scale	How to rate	Is it a behaviour scale? … I still think it should be separate [a separate scale]

Group 3, page 2

Scale 2	Scales	Structure	Again we were wondering about two scales here; the problem with activity levels, separate from joint and/or sustained attention ... Yeah, yeah, that really stands out doesn't it? There so much to take in.

Group 4, page 4

Scale 7	Scale 1 and 7	How to rate	I think this is hard to tease out from Scale 1 for infants [probably until they are talking, or at least the under-one's], because you're seeing the same thing …could fit in hyperarousal – where does that fit? …exaggerated startled responses, or floppiness – where does that go with emotional regulation, as well

			…its close, isn't it?

Group 4, page 5

Scale 7 and 4	Scale	How to rate	Doubling up with feeding: *At more severe levels, the infant may be underweight and presents with failure to thrive due to discomfort and distress of feeding. Problems with sensory regulation may manifest as emotional dysregulation or emergent developmental delays.*

Group 5, page 4

Scale 5	Scales	Where to rate	*Cognitive, motor* - I read that and thought: do they have a *motor* developmental delay? … Which I think scale 6 comes into play as well … So we're just rating language and communication in this one?

Group 5, page 6

Scale 9 versus 1	Global	Where to rate	Transposing emotion and related symptoms… You could almost plagiarise the headings in ICD 10, to put in this scale

Group 5, page 6

Scale 9 versus 1	Global	Where to rate	Back-arching - scale 1 – overlap

Group 5, page 6

Scale 10 and 2	Description	Clarity	Joint action better placed here rather than scale 2?

Group 5, page 7

Scale 11 and Scale 9	Scale	Where to rate	Self-care does not belong here. Should it be independent from environmental exploration – if important, then paragraph break, or separate Scale for environmental exploration [Scale 9] … Two quite separate areas

Group 5, page 7

Scale 13	Scales	Delete or split scales	Socialisation is elsewhere and Anxiety is elsewhere so you actually don't need it …so ditch it …and you could split one of the other two … Not have that one but an elaborate on one of the other ones – developmentally does not fit with age group [school and backpacks mentioned here]

Group 6, page 1

Scale 1 and 9	Scales	Structure	Include frozen …some of this you will pick up in [scale 9] [discussion re the confusion of over-controlled versus under-controlled and which scale] … Frozen watchfulness is in [scale] 9 …add: do not include – see scale 9 …over-controlled comes up more in [scale] 9 …make it éxternalising" versus "internalising" …we changed that but no-one knew what it meant, so we change it to over- and under- controlled

Group 6, page 2

Scale 4 and 9	Education and Training or instruction	Where to rate	If eating to regulate – is that Scale 9?

Group 6, page 4

Scales 9 and 10	Scales	Where to rate	It's got *indiscriminate and overfamiliar social interactions* in this scale as well – do we want it in this one and the one before? Now that we've just spent all that time adding it in to the one before? I think there is a difference between disinhibited to a social interaction… Disinhibited can be an emotional type response. Is indiscriminate different to disinhibited? … At risk of rating it twice

Group 6, page 4

Scale 5 and 10	Scales	Structure	Pragmatics here, rather than [Scale] 5 Understanding, expressing and functionally using language

Appendix 3.3 – Theme: Specialist Knowledge

Group 1, Page 1

P4 – Key Principles	Education and Training or instruction	Specialist knowledge	"I circled the one about *Ratings are informed by familiarity with, and a good understanding of, infant and child development* mainly because I think, well I suppose *familiarity* and *a good understanding of,* is not like, expert knowledge,...but that's the area I personally would like more, personally, more education or information about to feel…that I was doing this accurately…so like nearly a qualifier in that there'll be some areas where you feel like I'm competent to comment on, say *social reciprocity* where I'm not competent to comment on perhaps *the impact of or* maybe an early childhood nurse who specializes in sleep would be competent to comment on Scale 8's *Problems associated with sleep*, so like having some, you can't be expected to go across them all but knowing that aohww [unsure sound], I think that's what you're looking at, where you don't have the knowledge in some areas, so you're a little bit making a judgement call instead of having an informed clinical decision on some scales, unless you're a developmental paediatrician in this area…so as any professional who would be doing it, whatever your skills would be, so as the Clinical Psychologist who worked with naught to four, *Scale 10, Problems with Social Reciprocity*, I'd go 'Yep, I'm competent to be a bit more able to assess on that' whereas [name omitted for confidentiality], you gave an example this morning of after two years of age, I don't know what's normal and what's not and there's no way to qualify, I think that's an issue.

Group 1, Page 2

P4 – Key Principles	Education and Training or instruction	Specialist Knowledge	I suppose the question is: How much knowledge is assumed about things outside the expertise of the rater, like, so for example, there's one of the scales that surrounds feeding, not feeding, speech and things, where it says that it should be rated even if there's not been further assessment by a speech pathologist or paediatrician, so it is asking for your, if you have a level of concern I suppose, so I guess its just a matter of maybe making clear in those principles that its, yeah, what level of expertise are anticipated with people rating because its kind of quite clear from the initial introduction that its targeting people who work with infant mental

			health primarily, but I guess, yeah, that that's going to be a range of people.

Group 1, Page 3

P4 – Key Principles	Education and Training	Familiarity	And I suppose that does happen when you do, I mean I am only familiar with the adult one, but you do take information from all different sources, not only one, and when you do it regularly, you are sort of familiar, by habit -but there's the expertise of the person you're talking to, to be able to make the age-appropriate comments, so I might be able to make it while I'm observing it, or I might be able to ask the parent, but yeah, interesting, I don't know, I don't know whether that's harder, but that was my thought about it

Group 1, page 7

Scale 1	Education and Training or instruction	Specialist Knowledge	I found difficult to separate what I *know* when I was reading this, it was a little but difficult, but I could see, yeah OK, I get that, but it was difficult to know what it might be like for someone who *doesn't* get it, who then has to actually use the tool – a bit hard to un-know what you know and it's a bit hard to try and look at it from the perspective of somebody who might be trying to be using it and they're not trained

Group 1, page 17 in answer to overarching question – does the tool do what is intended?

Specialist knowledge	I think it's got the potential to document the progress of infants and young children and make overt changes, but it comes back to the wide range of changes that occur between 0-47 months. I think that the tool has the capacity to pick that up, if the rater has the capacity to pick that up. My concern is that the tool requires some statement around who's going to be able to use it the way it was intended, because if you look at the group who put it together– highly specialised clinical people, but then I wonder about someone else picking that up

Group 2, page 5

Scale 5	Education and	Specialist knowledge	I looked at this and thought I need a speech therapist in a way, to have a really good look at this

	training or instruction		area… People may think they've got a language problem when they've got an articulation problem [motor]

Group 2, page 6

Scales	Education and training or instruction	Specialist knowledge	I think there's learning effect – the more you do these, the more it sorts itself out; where to rate this, or where to rate that.

Group 2, page 6

Scales 14 and 15	Education and training or instructions	Specialist knowledge	My only comment is that service providers are often confused about the system and knowledge – it's almost a given that there be problems about knowledge. Translating that into accessing the system is also different – because people may have knowledge, but they may not; I thought that [Scale] 15 was really good – managing how to ask for assistance and also navigate them are a different skill

Group 4, page 5

General	Education and training or instruction	Specialist Knowledge	If this tool is to be used by many, we don't want to make it something that can be used by someone with a specialist point of view – that's a problem I think …OT to advise on wording around sensory issues [for the tool in general]

Group 4, page 5

Scale 10	Education and training or instruction	Specialist knowledge	Level of knowledge of the clinician [describing scenario where a child gives the clinician a hug at the beginning or end of the session and an inexperienced clinician may think that they are being social, but it could be a problem because it's a relief; not being social – need to recognise that; disorganised attachment, who do the indiscriminate social interaction]

Group 5, page 3

Scales	Global	Specialist knowledge	I doubt this will be used for children under two, in the present situation… The point is you've got to

			make it understandable by someone who doesn't very often use it

Group 5, page 6

Scale 10	Education and training or instruction	Specialist knowledge	What is normal eye gaze – infants do turn away to regulate, come back after regulating include the word *sustained* avoidant gaze

Group 6, page 1

Scale 1	Education and training or instruction	Specialised knowledge	The description is not going to point people in the direction enough, if they're not a specialised service… The reality is, across Australia, most people would be generalists [referring to infants]

Appendix 3.4 – Global Themes – General

Scales	General	Continuity	I really liked the fact that you actually looked at the context of the HoNOSCA.eg, they've got 4 other kids under five, the husband's violent, whatever it may be… even including it, whatever the content is, for me, it felt like a breakthrough, and I was very excited, so I loved it; really, really good, particularly at this age group…because it brings in that issue… about the trauma [its usually] only the dyad work, and that doesn't reflect the whole context

Group 1, page 4

P4 – Key Principles	How to complete	Which scale to use	...one thing I noticed that it said the scales should be rated in a certain order? And I'm interested in that's how the HoNOS and the HoNOSCA should also be done, so that should make it more uniform with those tools, OK. 'cause when I first went to look at it, I noticed the first one said now don't rate x,y and z in this scale, you rate that in that scale and I'm thinking, oh OK,

Group 1, page 4

P4 – Key Principles	Education and training	How to complete; order of scales Change order to scales	I agree. I had the same experience when I did a trial on a person, like in practice, but yeah, understanding the reason behind the instruction to rate one after the other and go from 1 to 15, if there's logic behind it, but I went through it and went oh, well, actually, and then I changed it around and I think that reflected better, I'm guessing was better

	Structure	Why in order?	…Yeah, 'cause I understood the whole measure, like I'm sure once you're really familiar with it, it would be different, but on the first time when I was doing my example kiddie, that's, like I had to go back; I don't know if the instruction is really that helpful if there is not like a logic to why you have to do it in order. …Why in order? and I wondered if it was about the uniformity of the tool, so making it more consistent with the other tools in the family of tools. …And it may be in the HoNOSCA research that there is a reason given around that and why you do it that way I… don't know – we've just always done it that way How compliant of you (laughing) I'm a good employee, aren't I?

Group 2, page 4

Scale 5	Global	Scales	Some issues are too broad and general, others too specific.

Group 2, page 4

Scales	Global	Scales	Some scales seem to cover a lot of symptoms; can be confusing as to what is being looked at - perhaps break into dot points to make them more understandable

Group 3, page 6

Scales	Global	Description	The infant, the parent and the relationship – those three perspectives are relevant to all scales

Group 3, page 6

Scales	Global	General	It's a lot better, I'm in the fact that we can even give specific micro-feedback speaks to the fact that the scale is a lot tighter… The one we tried 12 months ago was a lot more ambiguous then

Group 4, page 1

General	Education and training or instruction	Subdivide age range	Concern: The age range is so wide; looking at it with respect to a baby that five months old, it's really hard to score a baby that age according to some of these parameters because a lot of its written towards a baby that is much older with intentionality things

Group 4, page 2

General	Education and training or instruction	How to rate	[discussion not to rate according to cause – need to make clearer in preamble]

Group 4, page 4

Scale 6	Global	Subdivide age range	Small infants are much more susceptible to tiredness, those sorts of things than an older child, so you'd expect less of an ability to say definitively "this is a problem" …particularly with one-off – where you have an hour

Group 4, page 4

General	Global	Subdivide age range Developmental stages	See the Louis Marco's they were done with nursing staff over the whole week period and with a lot of 24/7 [much agreement heard]…exactly, and so when you're wanting people to do this and it will mostly be in an outpatient setting, I think its hard to assess with a baby …I agree …but its also very important

Group 4, page 5

General	Global	description	It would be useful to go to a student perspective and at the opposite end, a specialist one …get the medical students to do it, and the clinicians

Group 5, page 7

Scales	Scales	Clarity	[scales 14 and 15] I would be interested to know why these were included originally; how they got to be there

Global - Focus Group Responses to Index of Scales

Group 1, page 1

Global	p3 – Index of Scales	Index	"Pretty straight forward"

Group 3, page 1

Global	p3 – Index of Scales	Index	Similar to the HoNOS and HoNOSCA - I think that's probably good in a way, it's familiar

Group 4, page 1

Index of scales	Scales	How to rate	At that age, would all non-accidental self injury be [Scale] 9?

Group 4, page 2

Index of scales	General	Usefulness	Also at that age, the baby's behaviour changes rapidly, so I suppose in that way the HoNOSI does pick up what I want, because between this patient's admission and the discharge [case mentioned] her relationship with her baby changed a lot, and so the infant's behaviours towards her did change that way as a rating of what's happening, it is worthwhile, so they're are pros and cons I guess.

Group 5, page 1

Key principles	Key Principles	General	Clear if you've used HONOSCA – looks like a downwards revisions of that …Fairly self-explanatory

Group 6, page 1

Key principles	Key principles	General	Clear

Global - Focus Group Responses to Key Principles

Group 1, page 5

Key Principles	How to complete	Difficult to follow instruction	I found it tricky to follow the Key Principle of *Do not include information already rated in an earlier scale*. [Agreement heard on recording] Because the scales are so often reflecting similar constructs, so I don't, yeah, I think its tricky from an assessment perspective, what's, for instance like especially the Scale 5 around *Developmental delays* and then we reference development throughout the entire thing, like its integrated across nearly every scale apart from the-how to complete the other ones which are around family, so like *developmental issues* could be related to Scale 6 and *Problems of physical illness* could be related to scale 7, *Sensory Processing Information* could be related to Sleep and other developmental stages, so following that instruction was tricky for me to not relate them (other respondent:) or to not use information from one rating or to have them impact on what you're thinking about the current rating, so even if you were to try and separate it out, it would actually be quite difficult to separate it out because of the integrated nature of all of those signs together

Group 1, page 5

	How to complete	How to rate; not source of problem	I thought it was good the qualifier and you mention it also in the preamble I think, that, the presence of the clinically significant rating does not imply the source of the problem or the position for intervention – I think that's really nice, very relevant

Group 1, page 6

	Purpose	How informs assessment process	So I guess then too, in saying that, then what might be useful in terms of the tool going forward is some statement around how they would see this tool informing the assessment process within a suite of tools to help make that statement that its not necessarily identifying a locus or a point for intervention, its actually identifying something

			that's happening with the infant or having an impact on the infant's social and emotional well-being, but I think then it might be useful, particularly if its going to be really broadly available and used for people and what we know if it comes in, and its used like the HoNOSCA and HoNOS are used, then there's gonna be non-infant-mental-health trained people picking it up and expected to use it

Group 1, page 6

	Useful in practice	Usefulness	This would be really helpful, like when I was talking to *Brighter Futures* this week, to use the language to be really specific about what the issues were with this child, so in other services as well [agreement heard] And like the adult ones used for NGOs as well, its a good way to discuss specific issues

Group 2, page 1

Global	Rating period	Time frame	So the rating period's two weeks. I think that's a good framework because you need some of that time because some of these parents may think that they're worse than they are, or not as bad as they are in the rating period and because it's only two weeks that can mean that it's hard to measure really in a way that will necessarily indicate how much of a problem it really is

Group 2, page 1,2

Global	Rating period	Wording	Related to the timeframe, I think its scale 13 about attendance, socialisation and stuff - there is a comment that says if it's in a holiday period you can use the last two weeks, I'm wondering then should all of the scales be rated on the same two weeks using that scale as opposed to using say, the last two weeks for all the other items and then a different two weeks for the…[pause] because I think the two would influence each other potentially. Having just seen a six-year-old boy with autism this morning, where it makes a big difference, that was one of the things I wondered about with that. I mean it may not be as big of a deal with younger kids, but it's just a thought I had about would that materially affect other scales if you were using two different datasets to rate it.

			(other respondent:) a couple of things. If you make an exception for one measure, that's going to be difficult for people to hold that in mind, because you do so many of these. The other thing that I'd be worried about is if we made the exception for that HoNOSCA scale, should we do that for all the other HoNOSCA scales that we do? And are they going to at some point be, the data compared the HONOSCAs for a six-year-old, a two-year-old and a 10-year-old. If you make an exception for one, you have to do it for the other (first respondent:) my sense would be… to just do the whole scale the same; if you're in the second week of the school holidays and you're gonna do it on the last 2 weeks of school, you just do the whole scale on the same period of time, and it was confusing then [agreement heard] ... But I'm sort of ambiguous in that, and I wondered what people thought of that? (other respondent:) two options: One, it's a research study, does it make a difference? If you do the two weeks, not two weeks, and there is the big issue that my heart's fluttering about, if you change it for this, you actually change it for all of them, and [sigh] but its worth thinking about for future iterations of HoNOSCA (first respondent:) I mean is it easier to remove that ambiguous sentence?... In a sense we leave it at the discretion of the rater with the HONOSCA as well (other respondent:) which is what I think happens anyway, when you're doing it, we do for kids on school holidays, especially over January and stuff, and it kind of… you have knowledge of the family… sort of on the ground I think that's how it works. You just adapt it without sticking too strictly to the two weeks

Appendix 3.4.1 – Global subtheme: Structure

Group 3, page 2

Scale 3	Scale	Structure	It feels a bit like a forced scale in some ways, its sort of adapted obviously from the HoNOSCA. You do see occasions of self injury, probably more in toddlers, probably less scalable I think than this

Group 3, page 3

Scale 5	Scale	Structure	Trying to capture all of those things

			… We have whole teams to work that out

Group 4, page 1

Index of scales	Scales	Structure	Issue: overt displays of distress and shutdown behaviours go together and I think for babies in particular they're quite different things that you're dealing with… should separate them

Group 4, page 5

Scale 11	Scales	Structure	*Problems with separating from parents when the infant is attending structured socialisation settings* …I find that more of an emotional issue, its not about environmental exploration

Group 4, page 2

Scale 3	Scale	Structure	… That for me feels like its been put in because its in the HoNOS and the HONOCSI and it really isn't applicable, and there are some 3 or 4 year olds perhaps, but that would fit in another scale anyway, and it sh*ts me, that sorts of thing, and I think ethics committees would have a major issues if you were going to use that scale – the idea of raising that babies will self-harm …when she jumped of the height, she was looking for the ball, rather than injuries [laughter] …I really think it's a major issue

Group 4, page 4

Scale 5	Scale	Structure	Language and Communication in babies is different in other developmental delays and worthy of looking at in its own right…and other specific developmental delays [2 separate issues]

Group 6, page 1

Scale 1 and 9	Scales	Structure	Include under controlled emotional regulation …withdrawal comes into scale 9 …I think the context is also that it needs to match up with the HoNOSCA…but for us the struggle is how we bring the infant into this…our feeling is that it doesn't work so well

Appendix 3.4.2 – Global subtheme: Length

Group 1, Page 3

P4 – Key Principles	Global	Length	I'd be interested to know how long, people who put this tool together, thought it would take, to be completed?

Group 1, Page 3

P4 – Key Principles	Global	Length	Its very lengthy to read, but I actually found that really helpful, um, probably *because* I don't think I'm an expert in infant development and stuff, I just thought that it gave a lot of really good examples and that it helped me to complete it really well

Group 1, page 17 – duplicate – also came out in global overarching question:

	Global	Length	Initially it was clunky and cumbersome, but you'd be doing that day in and day out really. … quite long, but once they get used to it…
	Global	Length	The glossary is very comprehensive, but it doesn't necessarily address all of the nuances of this age range which is *such*…developmentally variable

Group 1, page 13

Scale 4	Length	Length	so that then goes to my previous question about thinking how long it would take me to use this tool

Group 2, page 6

Scale 13	Length	Length	On the verge of what would be considered long in the include box … The shorter, more succinct, the more practical it will be in day-to-day use

Group 3, page 2

Global	Length	Length	We're thinking to summarise it a bit, generally, if there could be less words…

Group 4, page 3

General	Global	Length - delete or integrate scales	[suggestion] – remove the whole scale- scale 3

Group 4, page 3

General	Global	Length - delete or integrate scales	In fact, in terms of working with mothers and babies, that has been the sort of feedback we have discussed here, that the whole thing is so big, its daunting, and so if we were looking to simplify it to 12 scales, maybe, you know, so that's partly what's been on my mind. There's some that we could put into each other, or remove. …and if there are those older kids who are doing things, they'll be covered in other scales

Group 4, page 3

General	Global	Length	That was my first thought when I first saw it; the size of it, and if practically people are going to use it, but its too big, it won't be used

Group 4, page 3

General	Global	Length – feasibility, practicality	Yeah. That is what happens in here. We've got long experience of using scales and then stopping using them because they're too complex. Do you know Stephen Matthey's scale? From Sydney? Its an adaptation; there's a scale by Anton Guedeney, a French scale which is 8 items and 4 points, so its 32 altogether and that was too much, 8 times 4, was too much for people to fill in – it's the Alarm Distress baby scale; and then Stephen Matthey who is in Sydney adapted it, its used worldwide, and he simplified it, its called the modified ADBB – Alarm and Distress de bébé and his is 5 points – pneumonic FEVER – Facial Expression, Eye Gaze, Vocalisation, Activity Level and Relationship to Examiner and his is only 10 points altogether as against 32, because its 5 parameters and you can only get 0,1 or 2, so we actually do that as a screen for all of our infants here, so we discuss it at Ward Round and because its 10 point; that's the maximum that they can get, its easy, and there's only the 5 parameters, so its actually working quiet well in here. We had a scale we developed ourselves long ago, which was mainly based on the mother's maternal health rather than about the infant, but it was about the relationship, and even though it was developed here (the Louis Macro), we abandoned it. The nursing staff were the ones who had to fill it in – that's it, it stopped. Even though we've got very good nursing staff and they're diligent, compliant, good people. That's a practical

			demonstration overall that FEVER's do-able, but its *minute* compared to this, its *minute!*

Group 4, page 5

Scale 7	Global	Length - delete	This is a tricky one I think. …scrap it? Yeah! [much agreement]

Group 4 page 6

General	Global	delete	Wonder whether it could be coalesced into something else? …I wonder if it can be scrapped? [Agreement heard]

Group 4, page 6

General	Global	delete	Because of all the excluding suggestions [Scales, 8,12,13], I think this is covered in those

Group 4, page 6

General	Global	delete	I think this is one of those that's filtered down from the HoNOS, then the HoNOSCA – I think that's one that isn't really relevant …if everything else is OK, this is going to be OK

Group 4, page 6

General	Global	delete	Scale 10 – social reciprocity – if the parent's doing that well, in terms of exploration and secure base, which is social reciprocity, then [Scale] 11's going to flow on, you could almost say it could be the behavioural manifestation of what's happening

Group 4, page 6

General	Global	delete	…and if its not [going to be OK], then it could be covered in other scales

Group 4, page 6

Scale 13	Global	Structure	This is another one that has been brought down from the HoNOS and the HoNOSCA – its about whether the child gets to go to child care and things like that – sometimes they're made to go even if it is a real problem being there, because

			they get taken there, even if attending is an issue – this is taking the school attendance idea - a 6 month old can't refuse to go to child care

Group 4, page 6

Scale 13	Global	Delete	I just actually think that should be dropped

Group 4, page 6

General	Global	Delete	Down to 11 [Scales]

Group 4, page 6

General	Global	Delete	Its of marginal value to this …I think scrap it! …it just adds to the burden of doing the scale if you have to do it every time and its not going to be relevant to the vast majority …and it doesn't matter if they don't go to things because they don't *have* to go to things, its not like a school aged child not going to school …kids don't have to go to child care …it could be included in Scale 15 *a lack of understanding about services or managing the infants' difficulties* – it could well be included - you could *easily* just say "Parent has difficulty getting the kid to child care when the child objects, not seeing the value of child care for the child, child needs limit setting" – you could *easily* incorporate it into [Scale] 15.

Group 4, page 7

General	Global	Delete	Its just feels like it was put in there to fit in with the others

Group 4, page 7

Scale 14 and 15	Scales	Delete/ consolidate	Problems with lack of useful information – could come into 14 – collapse them together

Group 4, page 7

Scale 14	Scales	Structure	14 and 15 are both there, but its not *adding* anything to the assessment, because one flows from the other

Group 5, page 1

Index of scales	Scale 1 and 9	Structure	This is the one that I had the most difficulty with, working out the crossover between Scales 1,2 and 9

Group 5, page 2

Scale 1 and 9	Scales	Structure	Scale 1 is about disruptive behaviour and Scale 9 is more about introverted sort of behaviours; depression, anxiety, phobia; I suppose that doesn't really take into account there might be underlying feelings of one set or another which one could regard as internal, but of course it manifests externally, but I don't know, I think that's making it far too complex and probably to just see those two scales as reflecting different aspects of behaviour problems – seems to me they do

Group5 page 2

Scales	Global	Clarity	Clarify that you're rating observation, not cause

Appendix 3.4.3 – Global subtheme: Consistency

Group 2, page 5

Scales	General	Consistency	Wondered if each scale is written by someone different; need to make it more consistent

Group 3, page 2

Global	Structure	Consistency	It looked like there might have been different people describing it and the language changed, so if there could be one editor using the same language all the way through

Group 5, page 5

Global	Scales	Consistency	It seems like certain specialists have contributed to the development of each of the scales… You get the sense there are slightly different styles, a little bit more detail in one and not the other… Ideally would have the same flow of language and similar level of detail for each scale

Appendix 3.4.4 – Global subtheme: Clarity

Group 3, page 4

Scales	General	Clarity	Language ambiguous at times – be more simple – use plain language eg putative.

Appendix 3.4.5 – Global subtheme: Sources of Information to use and how to rate

Comments related to Sources of Information (which need to be made clearer in the preamble)

Group 1, Page 2

P4 – Key Principles	Education and Training or instruction	Sources of Information	The other thing I, in terms of, I don't know if this is appropriate or not, but, I don't know about you guys, but when I looked at it, I thought: *How much of this am I meant to observe? How much of this am I meant to rely on parent reports?* And there's elements of it I think I'd actually have to go and observe the infant with other caregivers or in other environments, or talk to other environment care providers to see whether the behaviours are more global, all happening in that moment…have you already said this?[referring to participant having stepped out of the room to take a call], no, so I don't know, when I looked at the tool, I thought yes I could use it, I could see it, I could use it but how much of this is supposed to be parent self-report because I think the preamble talks about getting the information from as many different sources as you need (?),so (other respondent reads: *Use all available information in your rating to reflect your judgement based on all sources of information available to you*) [back to first respondent]…so that might be parent report, it might be care provider report, it might be what you observe, there's a whole lot of things and to actually answer some of these questions in a manner that I would think was actually accurate, I would probably do all of the above and so I'd be observing, I'd be asking, I'd be asking other care providers

Group 1, Page 3

P4 – Key Principles	How to complete	Sources of Information	I think your point about, and maybe this is more relevant to specific scales, but your point about the parents' view of the issue versus the clinician view, versus the view of any, you know those things sometimes are difficult to reconcile and so, um, yeah, kind of assimilating that information and having confidence in your judgement about rating

Group 1, Page 4

P4 – Key Principles	How to complete	Sources of Information	(another person:) and I'm just imagining trying to observe all of the circumstances that I'm being asked to rate

Group 1, Page 4

P4 – Key Principles	How to complete	Sources of Information	(another person) Well, I mean for example, Sleep. I mean you're not going to be sitting with a parent for example at 7 o'clock at night when they're trying to, you're relying on them to say this baby is impossible to get to sleep or you're relying in their expectation, so if their expectation is that their baby will sleep through the night, where actually developmentally that's not appropriate, you know, you need to have the skills to explore that with the parent in terms of *Is their expectation developmentally appropriate, or not?* Is there a mismatch there? Which then perhaps comes into the other scale around

Group 1, page 14

Scale 5	Education and Training or Instruction	Sources of information	I came back to the same thing about [rating 2] *mildly severe concerns noted across more than one setting and in comparison to peers of a similar age* – not an easy "do" ie not easy to collect that information. It's not just something you can collect on observation; you'd have to be talking to multiple care providers or settings.

Group 1, page 15

Scale 6	Education and training or instruction	Sources of information Clarity	[this scale] brings in the reports from others. Because this scale is overt in stating that in the inclusion, it made me think "does that mean I shouldn't include information from others in the other scales".

Group 1, page 15

Scale 10	Education and training or instruction	Sources of information	*The scale includes problems with emotional attunement and misalignment between the infant's needs and the parents' or carers' responses.* The only way that could be measured… Is to be observed.

Comments related to How to Rate:

Group 3, page 4

Scale 5	Education and training or instruction	How to rate	If you have a delayed infant, where is that recorded?

Group 4, page 4

Scale 6	Global	How to rate	Do the colic-y babies get mentioned in here? …in the feeding one [Scale]…it includes…gastrointestinal symptoms …it doesn't just fit in one [Scale] …seems cleaner in the adult one…more clear-cut

Group 4, page 5

Scale 9	Scale	How to rate	Acting out or shutting down should be two separate scales for this group rather than covered in both

Group 6, page 4

Global	Scales	Ratings	I prefer this type of rating where you obviously have to refer back to the explanation, rather than the explanation being embedded in the rating; I think that's easier …would make things easier …that's how the HoNOSCA's done

Appendix 3.4.6 – Global subtheme: Global Score

Group 1, Page 2

P4 – Key Principles	Rating	Global score	I guess what I'm not familiar with is, um, the document that I've downloaded, it doesn't actually show me what the outcome of the rating is, so (another person:) – You meant like a score out of… (original speaker:) – Yeah, at the end, yeah in terms of the, so the glossary doesn't like, um, and look I'm not really familiar, it's a long time since I've used the HoNOS and I've not used the HoNOSCA, so when I looked at it at the end, there's no sort of -overall cut-off score (2nd person again) -yeah, or something like that; I know we enter the points as separate items, but it doesn't give me an overall picture of it all

Group 1, page 5

	Global score	Global Score	(other respondent:) and that's why I think it is, and that's what would be helpful for me, if I was rating to understand, how the overall, what the overall rating means at the end, so I guess that its, I understand that its saying it's a measurement of the infant's mental health, I'm assuming… [Facilitator confirms: yes; social and emotional well-being] (Respondent:)…that's right, so its not an indicator for intervention (other respondent:) although they do say for all of these scales that a rating of 2 and above is clinically significant, which to me implies a need for some sort of intervention. If something's clinically significant then you can't rate it a 2 and then go Oh don't worry, you know, like, I think there's that sense of (other respondent:) but at the same time there's that sense of…it may not be the locus of intervention, with clinically significant it may not be what's driving the (back to previous respondent:) the intervention may not be with the infant the intervention may be what's driving it I think that's an important distinction [agreement heard] because otherwise you'd start addressing, say you start addressing all these say Scale 5,6, and 7 when the issue is really related to a bigger picture and this scale should not indicate what to do

Group 1, page 8

Key Principles	Global	Global Score	…that's what I'm talking about; it would be useful to talk about what other tools might assist and overall assessment like rating the quality of the relationship, parental mental health, you know, that sort of stuff as well

Appendix 3.5: Theme: How to Rate

Group 1, page 10

Scale 2	Education and	How to rate	I think I went to environmental contexts. I just said this could be very difficult to rate outside the caregiver context and then may require the

	Training or Instructions		assessor to try to interact with the infant, which, I do know that you guys, but I tend to think of not directly interacting with the infant but to observe the interaction…not sure whether this is now asking me to interact with the infant to see if I would get a different response from the infant to what I observe with their caregiver, so what I've said is this could be very difficult to rate outside of the caregiver context and then may require the assessor to try to interact with the infant to assess, that is assess with the caregiver and then assessment the clinician providing activity and appropriate support to see if the infant does things differently I took the multiple contexts to mean that it's more pervasive, so you might from the day care giver, is not limited to one personal situation
Scale 2	Rating	Interpretation	Rating 4 says there are severe levels of overactivity/underactivity which impacts negatively on the infant's capacity to engage and achieve developmental milestones across *multiple contexts*. What does that mean? Are we talking about caregiving contexts or life contexts or are we talking about multiple contexts of the infant?
Scale 2	Education and Training or Instructions	How to rate	I took the multiple contexts to mean that it's more pervasive, so you might from the day care giver, is not limited to one personal situation
Scale 2	Education and Training or Instructions	How to rate	So that means you've either got to go on watch it, or you've got to talk to people
Scale 2	Education and Training or Instructions	How to rate	So do we go with the majority then? … What I mean is the tension side of things usually becomes an issue with one particular person, usually the mother, and it may not come across in other people, so what are we measuring?
Scale 2	Education and Training or Instructions	How to rate	You're measuring the infant. So in that case you wouldn't rate 4; it's not sustained across *multiple contexts,* you'd rate 2 or 3, that would be the difference

Group 1, page 12-13

Scale 3	Rating	How to rate	The medical condition, does that include autism, that sort of thing? (Other respondent:) and should it give examples if it does…

			(Other respondent:) …you'd be rating it anyway… because it is dysfunctional (Other respondent:)…but when you have the condition, your capacity is on a different level (Other respondent:)…but then you're accounting for condition for everyone (Other respondent:)… That's what I don't understand in terms of this is a scale that measures the capacity, is that correct? (Other respondent:)… Does exclude secondary to medical condition but then on scale 6, it then doesn't talk about self injurious behaviour so theoretically you could have a three-year-old headbanging, that you would rate zero…because it was due to for example being on the spectrum but then you wouldn't rate that anywhere else, if you're excluding it here…if you're excluding it here, it probably needs to come in somewhere else
Scale 3 and 1	Rating	How to rate	… On Scale 3 is says <u>exclude</u> *Self-injurious behaviour secondary to a medical condition* but in Scale 1, I had to *include behaviour associated with any disorder* and so I wasn't sure why I was including it in scale 1 [behaviours, whether disruptive or self-injurious] and excluding it in the third Scale…and maybe I don't need to know – maybe I just need to do what I'm told (laughs)
Scale 3	Rating	How to rate	But it also does instruct you in the instructions not to consider the causes, so it is a little bit tricky – I understand when they are saying exclude because you're rating on another scale, that sits perfectly comfortably, but just saying *don't worry about it*, like I wouldn't feel comfortable rating a little kid headbanging on a zero just because he had a diagnosis of ASD
Scale 4	Rating	How to rate	To accurately assess this, I'd need to observe feeding, because so many of these things, if they're parent reported are much harder to rate for accuracy when the parent's reporting it, so I would need to observe it

Group 1, page 13

Scale 4	Rating	How to rate	If you're being objective and saying someone recognises that this is a problem, even if it is just the mother, and you're rating it based in that, then it is sort of flagging something that needs to be discussed with someone who can better assess it, which I wouldn't be doing for feeding behaviour, but you know quite often they have someone like an early childhood nurse or someone that can

			assess that, and I suppose if it is a tool to have a conversation with people, to identify problem areas, then its better to overrate than underrate, isn't it, even if it is just the mother that's concerned about it. Next time you rate it at 13 weeks, or whatever it is it might have been fully resolved and then that would reflect the new rating

Group 1, page 13-14

Scale 4	Education and Training or instruction	How to rate	My only thing with this and I don't know how you capture it, is that feeding, like sleep, the developmental stage is difficult, so if you are measuring this during the infant's transition to solids, say, then they could rate much higher, you know then you could have parents who are struggling with that…I don't have a solution but I think maybe there needs to be a bit of a recognition about rating that occur during times of transition (other respondent:)Maybe even in a general comment right at the beginning (first respondent:) yeah, like if the child is starting daycare while there's a major life transition, it's probably not the right time, or they're simply tired (other respondent:)You watch a baby feeding when its tired compared to a baby feeding when its awake and playful… absolutely there could be quite a difference - it comes back to how that tool is used by the rater, do they know, are they aware enough to say is this how it usually is? to the parent

Group 1, page 14

Scale 4	Education and Training or instruction	How to rate	I would be reluctant not to do a higher rating because they're at a transition phase though, because clients come in contact with us at transition phases because they are having problems, so excluding the fact they're in a transition would not be helpful because you're going to mix (other respondent:) no I didn't necessarily mean excluding, I just mean that… (other respondent:) sort of regardless of whether they're in a transitional phase or not, because the transitional phase will resolve itself next time you do the measure again

Group 1, page 15

Scale 7	Education and training or instruction	How to rate	This scale felt very difficult to rate. It may require a series of specific interactions to be able to observe and scale it, and how much would the scale respond to a parent being asked questions?

Group 1, page 15

Scale 7	Description	How to rate	Good that it's fleshed out as it may not be in the knowledge scope of a lot of people who'd be rating this, and it is important.

Group 2, page 2

Scale 1	Education and training or instruction	How to rate	With rating 3 and 4 it's important to consider the duration and I also wondered whether you excluded traumatic experience from the parents

Group 2, page 2

Scale 1	Education and training or instruction	How to rate	For the rating points down the bottom, some have less explanation, some have more, so it needs to be a bit more consistent for me understanding and what would *moderate* mean versus a *mild*. The moderate has got the explanation

Gourp 2, page 3

Scale 1	Additional tools	Assessing accurately	When parents are reporting this… There are tools… I was thinking about the duration… They can actually give them data to then be more specific around these scales because parents might think it's all the time, and in the record for a week, and it's actually three times a week for 10 minutes, but we feel like its all the time, therefore it comes out as a moderate or severe... That's where frequency and duration comes into it - parent informer as well; parents with mental illness often catastrophise the rating [Agreement heard]

Group 2, page 3

Scale 1	Glossary	Assessing accurately	(Other respondent:) that could go into the glossary is well, other things to include (Other respondent:) clinically significant tantrums versus normal tantrums (Other respondent:) he doesn't stop moving; he has ADHD!

Group 2, page 3

Scale 1	Education and Training or Instruction	How to rate Wording	I was unclear if I was going to rate in an increment, what you mean by disruptive behaviour in an infant. I don't know if that word is appropriate; an infant who's quite passive and listless the behaviours' of concern but its not necessarily disruptive… where you know they can't receive any care from their environment, where does that get coded, because that's quite significant? (other) overcontrolled emotion regulation …Yes, exactly. So where does that come in? …Scale 9

Group 2, page 3

Scale 2	Rating	How to rate	If you have two problems of moderately severe intensity, is that a reason to score it as a 4? Rather than 10/10 on one thing, can two 7's [mean he same thing]…making it explicit rather than implicit – I think people will probably do that anyway.

Group 2, page 6

Scale 11	Education and Training or instruction	How to rate	Do you consult with other people to complete it, or is it purely based on the clinician's observations? …Reflect caregiving level and the parents ability to recognise the child's development capability, so you do need observational data in this setting. … It's a very subjective rating. If its consistent over time, that's the important part.

Group 3, page 2

Scale 1	Education and Training or instruction	How to rate	What do we rate it on? I assume it's over the last two weeks, so parental descriptions will have a large weighting, whether you rated differently on the mental state observations and kind of like what you see

Group 3, page 5

Scales	Global	How to rate	The use of the scale in its entirety will have a lot more clinician prejudice than older kids

Group 4, page 5

Scale 8	Rating	How to rate	…everyone gets a 4

From Global Comments and related to how to rate:

Group 1, page 8

Scale 1 and 9	How to complete	How to rate	They actually do talk in Scale 9 about that sort of restriction and affect and withdrawal

Group 1, page 9

Scale 1 and 9	Description	Clarity	I find it a little bit like *Problems with disruptive behaviour /under controlled emotional regulation* I'm not clear what the topic of this scale is. Is it difficulties with emotional regulation, such as that they're expressed such as in disruptive behaviour, 'cause emotional regulation you could have the internalising kid, or the externalising kid and it kind of gives maybe overemphasis to disruptive behaviour which we already know that disruptive behaviour is disruptive behaviours that are gonna me more easily identified than the kids who internalise but they're still having problems with emotional regulation, so maybe the scale is meant to be emphasising that style of emotion dysregulation, but that's not like you're not seeing four month olds displaying disruptive behaviour for instance but you will see 1 month olds with poor emotion regulation and a tendancy to internalise or something, you would see that, so it might, you'll miss those ones (other respondent) well no you'll get them on scale 9, but I am wondering why we need, why those two things are being pulled apart 'cause Scale 9 is about emotion and related symptoms, so is Scale 1 really about behaviour?

Group 1, page 9

Scale 1 and 9	How to complete	How to rate	Scale 9 in the exclude box it says *Disruptive behaviours resulting from emotional distress* –

			rated at Scale 1, but the *emotion associated with the disruptive behaviour is rated at scale 9* – I don't know if that helps (other respondent) yeah that makes that clearer, but I'm not sure that it clears up, and because in fact they don't put exclude, like on Scale 1 maybe they need to talk about what you'd rate at Scale 9

Group 1, page 15

Scale 9	Scales	How to rate	The exclusion states "*physical sequelae of psychological disorders or medication – rated at scale 6*". I thought Scale 6 was focusing on physical, not psychological.

Group 1, page 15

Scale 1 and 9	Scales	How to rate	Behavioural versus emotional. Crossover between [Scales] 1 and 9.

Group 2, page 3

Scale 1 and 9	Scale 1 and 9	How to rate	There do seem some crossovers between 1 and 9

Group 2, page 3

Global	Glossary	How to rate	That may be something you put in the glossary, for example overcontrolled regulation be put in Scale 9.

Group 2, page 4

Global	Scales	Where to rate	*Self soothing behaviour* and some of the sensory seeking behaviour in another scale – double up …[Suggestion] where you code sensory seeking, refer back to where you code in Scale 3.

Group 2, page 5

Scale 5	Description	Where to rate	Motor planning could be covered under activity?

Group 4, page 2

Scale 3	Education and	How to rate	0-6 month old: at that age "distress"; might scratch. Have to rate as 9 [no problem/not applicable]. The issue is if a whole age group is

	training or instruction		rated as a 9, the scale is not going to work. The way it is, you can't do anything but give it a 9; or there's no problem, but there is a problem because the baby's distressed, but they're not self-harming.

Group 6, page 2

Global	Rating	Clarity	Generally I think the descriptions are adequate… Is to slowly move down to each category [ratings], they don't necessarily even reflect what's in the description …[Rating 4] *Overt inhibition of response to pain/discomfort and lack of self protection and self soothing* does not match the descriptor at the top.

Group 6, page 4

Global	Ratings	Clarity	Wording: Only bring in *severity* at rating 4 – make more consistent

Appendix 3.6: Clarity in relation to individual scale descriptions and/or ratings

In this Appendix, a copy of each Scale is provided for easy reference when reading the Focus Group responses, found below each Scale.

Scale 1: Problems with disruptive behaviour /under controlled emotional regulation	
Include	Oppositional & disruptive behaviours are normal in the development of young children as they explore and begin to individuate. This scale addresses problems with the age and developmentally appropriate capacity of the infant to manage strong feelings, without recourse to age inappropriate levels of overt disruptive behaviours. Clinically, the identification of age-inappropriate emotional regulation does not indicate the source of any difficulties. It may be expected, though not invariable, that regulating emotions connected to hunger, tiredness, and separation may be more prominent with younger infants while overt aggression or rage may be prominent with the older children. Include behaviour associated with any disorder (such as hyperkinetic disorder, depression, autism). Include the capacity to manage intense feelings of hunger, tiredness or separation from the primary caregiver. Include difficulty settling, demanding, whining, undue irritability, excessive crying, frequently arching back and stiffening coupled with turning away from all eye contact, inhibition of expected interactive response, physiological indicators of stress (hiccups, yawns, non-injurious scratching). Include physical or verbal aggression (e.g. pushing, hitting, biting, kicking, teasing), to others (e.g. children, parents or other caregivers, siblings, familiar adults or strangers), animals or objects (e.g. toys). Include oppositional behaviour (e.g. defiance, opposition to authority or tantrums).
Exclude	Problems associated with feeding and sleeping scored at scale 4 (feeding) and scale 8 (sleeping). Problems directly associated with physical health illnesses or disability scored at scale 6. Problems associated with self-injury rated at scale 3.

Rating	**Description**
0	No problems of this kind during the period rated.
1	Minor problems that are related to a transitional developmental stage, illness or circumscribed change in family circumstances.
2	Mild but definite disruptive behaviour, likely to be limited to one context.

3	Moderately severe disruptive or aggressive behaviour, likely to be in more than one context (e.g. home, childcare, other social situations).
4	Disruptive or aggressive in almost all activities At least one serious physical attack on others or animals, or serious destruction of property.

Scale 1: Focus Group Responses

Group 1, page 6

Scale 1	Wording	Clarity	I had a particular issue with the word *settling* in here because it deliberately excludes sleep and I think that lots of people, with infants, when they talk about settling, what they talk about is settling them to sleep, which is actually addressed in a different scale, so you may want to look at another way of expressing because what I think they mean is being able to calm down and you know, regulate, um, rather than rocking them until they go to sleep, which is the commonly used way that parents will talk about settling, and that's quite specifically addressed somewhere else, whereas I would think about a different way of expressing that -calming down or -yeah, absolutely -yeah, just those specific little… -and when you read number 8 it says settling as one of the includes

Group 1, page 8

Scale 1	Title	Scale	(other respondent:) but that is actually inhibiting, that's actually *overcontrolling or inhibiting* rather than under controlling so maybe the name of the scale could be reconsidered [agreement heard]: because then that would fit as inhibition would fit as if it was difficulty accessing help in managing their emotions (other respondent:) or they're not safe, they don't feel safe…and this is where I need to come back to think about what I am rating 'cause I don't want to start getting into interaction type stuff where you see, where you're rating the quality of the relationship and that's what I keep reminding myself of in doing this tools is I'm actually not, in most of the scales, rating the quality of the relationship

Group 1, page 8

Scale 1	Description	Clarity	Need more clarity about disruptive behaviour

Group 1, page 9

Scale 1	Title	Clarity	I am struggling to figure out what it is asking for, to measure…it seems like its saying lots of things, trying to put into one thing

Group 1, page 9

Scale 1	Title	Clarity	Or change Scale 1 title to reflect if its about behaviour then *Behavioural problems following* issues with emotion regulation instead of problems with both emotions and behaviour

Group 1, page 9

Scale 1	Title	How to rate	Because it feelings like that box on Scale 9, the exclude box, that this is expecting you to rate behaviour and this is expecting you to rate emotion, but I don't think that's necessarily clear in Scale 1

Group 2, page 2

Scale 1	Wording	Clarity	With the diagnoses: depression, autism, one of the difficulties I had [referring to a 24 month old with ASD], does that have a rating of two or above because he already had a diagnosis, but the diagnosis is quite mild and wasn't given by our service so there is also a bit of disagreement around it, so I found that a little bit tricky My sense of HONOSCA historically is that you try to put diagnoses aside and just think about broadly, problem areas. So it doesn't actually matter whether the child is depressed or has an intellectual disability on the spectrum or how bad is it. It's just isn't there, or is it not there and you leave the speculation for other stuff.

			That's how I've done the HoNOSCA. I don't know if the HoNOSCA scaling actually talks about diagnoses… made me a little bit confused, probably if it wasn't there I wouldn't have thought about it

Group 3, page 1

Scale 1	Scale	General	Got our tick of approval

Group 3, page 1

Scale 1	Wording	Clarity Specialist Knowledge	Title – the words *disruptive under control emotional regulation* - I wonder possibly about the word *dysregulated* might be more appropriate for infant clinicians, but …not as clear for people who aren't; haven't worked in that field. The word disruptive or challenging, they tend to me, to be developmentally appropriate for older kids and I wonder whether it could be added

Group 3, page 1

Scale 1	Ratings	How to rate Title	The severity ratings then, it's become synonymous with aggression … Would it be a better way of titling it as an aggression scale … Yeah, because dysregulation is under sleep and feeding and other sorts of scales … Then how would we describe aggression in a six month old … Make it clear that this scale is about aggression… I think the word disruptive is…Is it an aggression scale? If it is then we should call it an aggression scale.

Group 3, page 1

Scale 1	Clarify	Ratings	Might want to include toddler defiance which isn't in that group …excessive tantruming ... Aggressive and defiant behaviours, which is sort of alluded to in the text but not so much in the descriptors down the bottom [ratings]… You might put some defiance to request by parents but is not pervasive and parent can convince the child to do things at some of the time, can overcome these with persistence

Group 4, page 2

Scale 1	Title	Clarity	Take out disruptive behaviour and just say "problems with disruptive behaviour"; it feels wrong to say that a baby is being disruptive, and I don't think we need to; an infant couldn't score a 4 on that scale, I imagine [agreement] [or]even a 3 [baby could not score]

Group 4, page 2

Scale 1	Rating	Description	Hiccup and yawning could be misinterpreted differently – "yeah they yawned, there's a problem"

Group 4, page 2

Scale 1	Rating	Description	[point raised by clinician to keep a six month old infant in mind as we go through the scales] "Turning away from all eye contact" [sounds made, indicating this is not appropriate]

Group 5, page 1

Scale 1	Education and Training or instruction	How to rate	This is the one that I had the most difficulty with in working out the crossover between this one and number 7. I thought there was some crossover between Scale 1,2 and Scale 9 Scale 9 talks about exaggerated positive and negative emotional response and some of the examples given in Scale 1 – *demanding, irritability* - I think it might be difficult to know which ones, where to score some of those things.

Group 6, page 1

Scale 1	Description	Clarity	*Under-controlled emotional regulation* is not in the description at all; it's all about aggression, which is an issue when we're looking at the little babies who generally aren't aggressive

Group 6, page 1

Scale 1	Scales	Structure	The little ones who aren't settled – we're not getting that into the picture

Group 6, page 1

Scale 1	Scales	Structure	It's feeding, sleeping and socialisation that can be impacted by being under-regulated or externalising as an infant

Group 6, page 1

Scale 1	Rating	How to rate	All about aggression/disruptive externalising

Group 6, page 1

Scale 1	Rating	Clarity	Age range is not reflected in the ratings …what this means in an infant …generic CAMHS are more likely to underrate the infants because they're not displaying aggression, so infants are going to come out underscored

Scale 2: Problems with activity levels, joint and/or sustained attention	
Include	Include problems with overactivity/underactivity, joint and sustained attention associated with any cause, including related to aspects of the caregiving environment (e.g. lack of appropriate stimulation, opportunities for motor development). Include problems with restlessness, fidgeting, distractibility, jerkiness or concentration due to any cause, including depression. Include issues of sustained as well as joint attention. Activity and attention difficulties may manifest in altered levels of vigilance, impaired turn taking in behavioural interactions, pronounced startle reflexes and rigidity. Where two factors appear to negate each other (e.g. joint attention problematic but sustained attention highly powerful), rate the most severe occurrence.

Rating	Description
0	No concerns during reporting period.
1	Minor periods of overactivity/underactivity.
2	Some vulnerability in joint and/or sustained attention however the infant's development appears unaffected. Mild overactivity/underactivity or restlessness but with age-appropriate support/structure, the infant can modify their activity levels.
3	Regular significant issues, at times, with overactivity/underactivity that is difficult to control even with appropriate supports. Moderate problems in joint or sustained attention that is typically unable to be modified.
4	Severe levels of overactivity/underactivity impacts negatively on the infant's capacity to engage and achieve developmental milestones across multiple contexts. Consistent and severe limitations in sustained and/or joint attention.

Scale 2: Focus Group Responses

Group 1, page 9

Scale 2	Education and Training or Instruction	How to rate	There's a statement that says where two factors appear to negate each other, rate the most severe occurrence, I found that unclear and difficult to understand what I was meant to do

Group 1, page 10

Scale 2	Education and Training or Instruction	How to rate	There's a statement that says where two factors appear to negate each other, rate the most severe occurrence, I found that unclear and difficult to understand what I was meant to do

Group 1, page 10

Scale 2	Description	Clarity	Rating 3 – *appropriate supports* - I am not sure what is meant by appropriate supports; is this referring to parental supports, or what sort of supports are appropriate?

Group 1, page 10

Scale 2	Description	Clarity	I personally find *appropriate* a difficult word because that can mean different things

Group 1, page 10

Scale 2	Rating	Interpretation	Rating 4 says there are severe levels of overactivity/underactivity which impacts negatively on the infant's capacity to engage and achieve developmental milestones across *multiple contexts*. What does that mean? Are we talking about caregiving contexts or life contexts or are we talking about multiple contexts of the infant?

Group 1, page 11

Scale 2	Wording	Clarity	I think that there is a little bit of repetition in the description *associated with any cause, including…* then in the next paragraph you say again: *due to any cause including depression*… Instead you could use some text to outline some examples of infant difficulties with joint or sustained attention, what that would look like, 'cause if a person is not used to how long should a baby be looking at Mum for, is that normalising, I think some breaks in eye contact but actually saying there is a noticeable problem with the child

Group 2, page 3

Scale 2	Description	Clarity	This might be where listlessness might be noted as well; gives a different emphasis.

Group 3, page 2

Scale 2	Wording	Clarity	Swap around the sentences on rating two, it just flows with joint attention is the second point in all of them, except in number to where it's the first point that's made

Group 3, page 2

Scale 2	Wording	Clarity	Joint and/or sustained – make consistent

Group 3, page 4

Scale 8	Education and training or instruction	How to rate	Also begs the question where you got our family who's really distressed, or even one family member maybe a mother who's really distressed, dad's not distressed, clinicians saying well actually this kid looks well rested, will actually its getting a reasonable amount of sleep; how do you rate that then – do you rated on what you're seeing with the baby, what you're seeing with mum you know tearing her hair out… Maybe you take these factors into account and coalesce them into some sort of mid range type rating?

Group 4, page 2

Scale 2	Education and training or instruction	How to rate	I don't think *startled reflexes* is a sign of activity

Group 4, page 2

Scale 2	Scale	Description	Is underactivity written enough? - a child who is under-compliant, very good, I don't know whether that's fully in that - Not as much emphasis on the underactive, the over-compliant, overbright, good sort of child - to pick up dysregulation from that level

Group 4, page 2

Scale 2	Scale	Description	Is underactivity written enough? - a child who is under-compliant, very good, I don't know whether that's fully in that - Not as much emphasis on the underactive, the over-compliant, overbright, good sort of child - to pick up dysregulation from that level

Group 5, page 2

Scale 2	Title	Clarity	*Problems with joint attention* and then later on Scale 10 is *Problems with Social Reciprocity*, which in part of social reciprocity is having joint attention and I wondered how useful the word *joint* was? …confusing…made me think of developmental issues …confusing …if you took out the word joint, it would be fine…*sustained attention and care-giving relationships,* or something

Group 5, page 2

Scale 2	Description	Clarity	Exclusion – need restlessness, jerkiness as they might be having medical difficulties …that applies across the board to all the scales…even in very young prems, you have no idea what the impact of that is really

Group 5, page 2

Scale 2	Description	Clarity	Physical illness or disability – few overlaps there … Either make it an exclusion, or you need to make it more explicit, that you include

Group 5, page 2

Scale 2	Description	Clarity	Is it model of attention? For example, Mirsky's [model] is famous and endorsed and it focuses on attention, selective attention, divided attention – if you want to go down to problems of attention, how neuropsychological do you go? Add descriptors. …could be more of a red herring…there is more than joint and sustained attention in those models – for people who are not familiar with those models, we were talking about autism, social reciprocity … *Joint* is not a term I am familiar with

Group 6, page 1

Scale 2	Scale	General	Easier than Scale 1 – statements are more appropriate across age group

Group 6, page 1

Scale 2	Scale	General	Happy with [Scale] 2 …its about scoring the infant, not the relationship

Scale 3 Non accidental self-injury	
Rating	**Description**
Include	With infants and pre-schoolers, the question of intentionality is less clear than with older children. While intention should be considered, it will not always be apparent and the clinician may draw on clinical experience to infer intentionality. Behaviours included here are essentially those that result in self-harm that are not the consequence of an accident. However, self-injurious behaviours and actions are rated here irrespective of any indication of intent. Self-soothing behaviour that results in injury or harm e.g. hitting, biting, hair pulling, head banging, rocking, cutting, scratching, excessive sucking leaving marks. May include lack of self-protective reflexes, inhibition of pain and reassurance responses e.g. when an infant is clearly hurt yet inhibits a response where other infants of the same age would be expected to cry, flinch and look to parent(s) for reassurance. Include pretend play involving self-injury such as running into hard objects deliberately in the context of play, making dolls hit their heads during play. Include attempts to stab self with a pen or other non-lethal object, cutting self with knives or scissors, deliberately jumping from a height with injurious intent, frequently discussing intent to self-injure.
Exclude	Self-injurious behaviour secondary to a medical condition. Accidental self-injury unless clearly from a lack of self-protective reflexes.

Rating	Description
0	No problems in the rating period.
1	Occasional occurrence of minor lack of self-protective reflexes, thoughts about self-injury, pretend play involving self-injury, deliberate physical recklessness.
2	Occasional actual self-injury including actions such as rubbing, scratching or rocking which lead to mild levels of physical injury. Pretend play that regularly involves self-injury or self-soothing behaviours that result in harm. Occasional episodes where self-protective reflex is inhibited.
3	Moderately severe occasions of potentially or actually physically significant self-injury. Preoccupation with self-injury. Repeated episodes of self injurious behaviour. In younger infants – inhibition of response to pain/discomfort and lack of self-protection and self-soothing.
4	Severe or regular self-injury occurs. Episodes of physically significant self-injury. Overt inhibition of response to pain/discomfort and lack of self-protection and self-soothing.

Scale 3: Focus Group Responses

Group 1, page 11

Scale 3	Wording	Clarity	This one is tricky saying *self injury* because I'm excluding baby infant-infants straight away... and then as I read further, I thought oh, its also a little bit of that self-soothing, injury is joint under this box, maybe it's just semantics, but I usually went "Oh! That is never applied to my under-1's" and then as a read more I thought "Oh, under this scale it would, um [unsure sound]"… probably just that word *self injury*
Scale 3	Wording	Clarity	And then it kind of relates to *intention*, which I think you guys are trying to flesh out in those lines, you've been talking about intentionality with a baby it's pretty hard anyway

Group 1, page 11

Scale 3	Education and Training or instruction	Reference to developmental stages	And the transition from the baby to the one that deliberately harms himself, maybe my two-year-old who is deliberately being himself on the head with something doesn't know what that's about but that can escalate into an older, deliberate thing, it's not clear, again, about the stages, what sort of happens at each stage

Group 1, page 11

Scale 3	Clarity	Clarity: Interpretation	*Lack of self protective reflexes* - is that kind of a different thing to the child performing self injury, by not protecting myself, it also means I'm self injuring myself? If I was not to protect myself as an adult, I could get that that was about self injury

Group 1, page 11

Scale 3	Title	Clarity: Interpretation	So even going to *non-accidental self injury/lack of self protective behaviours* as a name in the scale [Agreement heard]

			I think that's good 'cause I think that name [referring to existing] will trip people up… (Other respondent:) I was thinking you are more trying to get at dysfunctional self soothing behaviour rather than the child trying to intentionally harm itself, but maybe I'm misunderstanding it Facilitator: and do you mean that across the range, or for the younger Respondent: Across the range, because even a 4 year old who might hit themselves on the head, in some ways its trying to get their needs met… It's a way of self soothing that's not working, so I suppose that comes from the way you see people's behaviour, if you see the behaviour is always trying to get their needs met

Group 2, page 3

Scale 3	Description	Clarity	Re biting – I have seen kids biting to self soothe themselves but they don't necessarily bite others to self-soothe –that was a question as to what that second statement was about?

Group 2, page 3

Scale 3	Description	Clarity	The term self harm in infants bothers me. The term self harm in adolescents has a whole different connotation [suggestion: use another word – self-injury?] … The connotation is slightly different. We don't tend to refer to adolescents as self injuring. …be consistent with title.

Group 2, page 4

Scale 3	Description	Clarity	*Non-intentional accident* needs to be defined

Group 2, page 4

Scale 3	Description	Clarity	*self protective reflexes* needs to be defined

Group 2, page 4

Scale 3	Description	Clarity	*Lack of safety awareness* eg jump out of a tree because they're not judging the distance or thinking of the consequences [include]. Non-intentional accident.

Group 3, page 2

Scale 3	Wording	Clarity; bad example	We were a bit concerned about having them making dolls hit their heads during play; that can happen with a kid who's got nothing outside, to be concerned about, not self injury, just part of play ... It feels like it should be almost a more binary type of rating in a way rather than trying to kind of tease it out into different components. If you have kids who are headbanging or slamming their heads into a wall or something, I think that is worth rating, but is it possible to actually tease out the little bits in between, it feels a bit forced - to fit a model rather than a scenario … We wouldn't be thinking about self injury at all for kids who are banging their dolls together

Group 3, page 3

Scale 3	Description	Clarity	Lack of self protective reflexes… Would be used under unusual circumstances

Group 3, page 3

Scale 4	Scale	How to rate	Does it include the relational aspect – rate rather than exclude relational

Group 3, page 3

Scale 4	Title	Clarity	Changed to "Problems with feeding"

Group 5, page 3

Scale 3	Description	Clarity	You could perhaps bypass the *attention* issue by using descriptors such as *concentration* and *focus* instead

Group 5, page 3

Scale 3	Description	Clarity	Clear. I thought its strength was that it's really clear to know what to include. Eg. don't worry about intentionality – I think that's useful. …its well done

Group 6, page 2

Scale 3	Description	Clarity	Category that is very difficult with the little ones that we would see… It's hard to keep the real tiny ones in mind when you read this

Group 6, page 2

Scale 3	Description	Clarity	Self-soothing behaviour – need examples for infant

Group 6, page 2

Scale 3	Description	Clarity	*Making dolls hits their heads during playing* I'm wondering if that's more kind of fantasy, rather than self injury? You could say rough play, not self-injury [discussion on what self injury means in infants] … I don't know if this actually applies to babies, the way it sits here – when you've got the feeding one …12 months and under, struggled to see how it would fit [for under 12 months]…Excessive sucking, or skin scratching or picking [might fit], or throwing themselves backwards…would not happen often …0-6 month old – we may not see it at all …include for 12 months and under – *observations might include…*

Scale 4: Problems with feeding behaviour	
Include	Feeding behaviours progress with development. The acknowledgement of problems in this area will be influenced by the duration, distress and incongruence of the concerning behaviours with the infant's age and age appropriate development. All glossary examples should be developmentally referenced. Include problems related to difficulties with breast feeding, bottle feeding and solids. Include all feeding difficulties irrespective of potential cause or solution. Include behaviours such as reluctance, resistance or refusing to feed; tiring or sleeping readily during feeding; feeding related distress (e.g. fussiness or crying); maintaining adequate nutrition which may result in nasogastric / gastrostomy tube feedings; sensory adversity; vomiting and difficulty in achieving developmentally appropriate food or feeding skills e.g. limited diet, consistent refusal of certain foods, groups, or types (e.g. solids), or modes of eating; little recognition of the relationship between hunger, feeding and satiety. Include feeding problems related to prematurity, physiological problems and gastrointestinal symptoms.

Rating	**Description**
0	No problems of this kind during the period rated.
1	Slight occasional problems only. Problems may be transient though typical for the infant's developmental stage.
2	Mild but definite problem with non-organic somatic symptoms. May require additional support to feed. Nutritional intake and growth are likely to be within expected parameters.
3	Moderately severe feeding problems. Some risk of nutritional problems.
4	Severe problems with feeding. May include a high probability of severe nutritional problems.

Scale 4: Focus Group Responses

Group 1, page 14

Scale 4	Wording	Clarity	What does *Mild but definite problem with non-organic somatic symptoms* mean? [Rating 2]…because if its not clear to me, it might not be clear to someone else

Group 1, page 14

Scale 4	Wording	Clarity	*Moderately severe* – is it meant to be moderate?

Group 2, page 4

Scale 4	Scale	General	Straight-forward.

Group 2, page 4

Scale 4	Description	Clarity	Description about what to include is quite clear

Group 2, page 4

Scale 4	Description	Clarity	Thinking about those early days of breastfeeding versus bottle, the parent can actually be quite defensive as some hospitals promote breast feeding and some don't, and then there's a whole movement outside of CYMHS of breastfeeding versus bottle and that may affect how parents may look at feeding in those really young infants. The scale is clear but the political overlay in that context is going to affect those parents and this question. … Parents don't answer the question. It's a subjective assessment by the clinician. The clinician will be able to assess regardless of the political connotation.

Group 3, page 3

Scale 4	General	Scale	Important scale to have – see a lot of …absolutely

Group 3, page 3

Scale 4	Description	Rating	Rating descriptors could be tweaked a bit…additional support to feed – could be

			interpreted with naso-gastric feeds…dump those examples and let the clinical decide

Group 3, page 3

Scale 4	Description	Clarity	Include: independent eating for older child

Group 3, page 3

Scale 4	Description	Clarity	Might get a kid who is feeding OK [nutrition and observed] and is all right, but you get a parent is stressed out of their head – we would you rate parent anxiety about feeding in their mind? Rate as 4 or 1 or 0?

Group 3, page 3

Scale 4	Wording	Clarity	… You make it then problems with feeding, rather than problems with feeding behaviour?

Group 3, page 3

Scale 4	Description	Clarity	Add in descriptor parental concerns regarding feeding, otherwise it won't be captured if there's a perfectly well infant

Group 3, page 4

Scale 8 and 4	Description	Clarity	Like that [it] accounts for parent distress in the family and I would like to borrow that in feeding as well; same perception issue in those two regulatory domains

Group 4, page 4

Scale 4	Scale	Description	Look at eating things they are not meant to eat; ingesting things other than food eg excrement or chalk

Group 4, page 4

Scale 4	Scale	Description	And also overeating isn't there and that can be an issue too

Group 5, page 3

Scale 4	Description	Clarity	What's *All glossary examples should be developmentally referenced?* … That's my question too. … Global comment? Specific to feeding?

Group 5, page 3

Scale 4	Description	Clarity	While I think it's important to include feeding problems related to physiological problems, sometimes there's physiological problems that can be changed if there is an underlying psychological issue as well compounding that, but sometimes it doesn't matter what you do, there's not much change - if we're looking at this as a measure to see how change is unfolding over time, some things won't ever change. … Exclude children where there is an organic cause due to feeding difficulties that are unlikely to change due to the congenital nature of the problem.

Group 5, page 3

Scale 4	Rating	Clarity	[Rating 3] *some risk of nutritional problems* and [Rating 4] … *A high probability of severe nutritional problems* do these need to be here? … Even a fussy eater may have nutritional problems … Does it include the obese infants? Does not seem to refer to obesity at all, which surprises me, given the current emphasis. They lay down fat cells and never lose them, so prognosticly it's a bad thing. … I would have thought it should be included … It's got the mental relationship between hunger, feeding and saiety … Include under-nutrition *and* over-nutrition

Group 5, page 3

Scale 4	Title	Clarity	*Problems with feeding* and eating *behaviour* to signify it does not just refer to babies

Group 6, page 2

Scale 4	Description	Clarity	*little recognition of the relationship between hunger, feeding and satiety* – may be a biological predisposition to over-eating, might be nothing

			to do with the relationship

Group 6, page 2

Scale 4	Description	Clarity	Need a sentence to cover under- as well as over-feeding

Scale 5: Problems with language or communication skills or specific developmental delays.	
Include	Include cognitive, motor, language, and communication delays irrespective of putative cause or solution. Problems in speech or language associated with any disorder or problem, such as a specific developmental learning problem. Concerns should be rated irrespective of whether additional professional assessment or intervention has occurred (e.g. paediatrics, speech pathology). Language or communication difficulties may include articulation, pragmatics, gestures, vocal quality or range, echoing, interference with vocalisation (e.g. dummy, fingers). Difficulties in this area may impact on ability to communicate and form relationships with similar aged peers and adults, emotional dysregulation, and behavioural issues in the context of difficulty in communication.
Exclude	Physical illness or disability problems such as vision and hearing problems (rated at scale 6). Do not include infants with global learning disability (Intellectual Disability) unless their functioning is below the expected level for them. There will be periods in the very young where the presence of a global learning disability is not conclusively established. It may be expected that problems may be rated on this scale that will subsequently be no longer rated.

Rating	**Description**
0	No language, communication or specific developmental delay issues identified during the rating period.
1	Some minor concerns noted by parent or clinician in regards to delay in speech and language development or developmental delays that may be monitored but likely to be within the normal range of development.
2	Mildly severe concerns noted across more than one setting and in comparison to peers of a similar age.
3	Moderately severe and/ or multiple concerns noted across settings compared with peers of a similar age.
4	Severe problems of language, communication or development exist. Likely to cause significant distress for the infant and/or family. Severe delays compared to peers of a similar age.

Scale 5: Focus Group Responses

Group 1, page 14

Scale 5	Wording	Clarity	There was one that said *mildly severe*, which I thought was a really odd thing to say, can't remember which scale it was in [Scale 5]

Group 1, page 14

Scale 5	Wording	Clarity	[Rating 2] *mildly severe*

Group 1, page 14

Scale 5	Title	Clarity	Referring to the younger population; communication skills: on this scale are you trying to assess communication skills or just a verbal communication skills? [Participant gives example of a three-month-old and makes a noise and contrasts this against eye contact]. [The suggestion is that "communication skills" may be too broad if we talking about verbal communication skills. Distinguish between global and verbal communication skills].

Group 1, page 15

Scale 5	Description	Clarity	There are not many examples on the younger age

Group 2, page 4

Scale 5	Description	Clarity	What do we do with kids who have a global delay?

Group 2, page 4

Scale 5	Title	Clarity	Put motor in the title … The first time I read it, I thought it was just on language and communication. … Helpful to have a paragraph or a sentence detailing what the motor skills are. ? Fine, ? gross ?Coordination ?balance

Group 2, page 4,5

Scale 5	Title	Clarity	Rename: Problems with communication, motor or specific developmental delays - because language, I would argue, would fall under the rubric of

			communication - expand on what to include/exclude

Group 2, page 5

Scale 5	Description	Clarity	Re: exclusion of intellectual disability. Hard to determine – suggestion – take that sentence out

Group 2, page 5

Scale 5	Description	Clarity	It's good how you've got the include and exclude… Needs tweaking, but I like it.

Group 3 page 3

Scale 5	Description	Clarity	[Related comments from speech pathologists who were not present, but had passed on their input:] Description does not include comprehension, sentence structure, vocabulary as a language or communication difficulty

Group 3, page 3

Scale 5	Description	Clarity	Also seems weighted towards communication, even though it's meant to cover all of the developmental domains…

Group 3, page 3

Scale 5	Description	Clarity	Have them separated to say language difficulties cognitive and problem-solving, which most difficult is include, it doesn't include toilet difficulties (in self-care)… There is quite a bit about language because it is appropriate; the language is the most obvious which brings kids to the attention of services whereas motor and cognitive difficulties are more subtle, so we need more examples of those, just like with language, so people who aren't as clear, it will be clinically useful for someone screening cognitive and motor… add a sentence for each of those

Group 3, page 4

Scale 5	Description	Clarity	Exclusion for global development delay is confusing. This in the development delay item, but then it excludes developmental delay; that's just cognitive version of a specific developmental delay. These distinctions become less meaningful as you move into infancy… Once you get to 6 or seven, it

			evens out… Take out the exclusion of developmental disability.

Group 3, page 4

Scale 5	Title	Clarity	... Take out the exclusion. … or just take out the "specific" … And I'd say " language, cognitive and motor skills"

Group 4, page 1

Scale 5	Scales	How to rate	I wonder for infants in particular whether the language should be in a category of its own separate from other specific developmental delays… we see specific issues where the babies aren't verbalising; quite different from the other developmental delays, and for them all to be locked in together…

Group 5, page 3

Scale 5	Scale	General	Broad. What to consider here.

Group 5, page 3

Scale 5	Wording	Clarity	*Pragmatics* – what does that mean? … Quite sophisticated some of the language that's used. Use less jargon, give an example.

Group 5, page 3

Scale 5	Description	Clarity	Why have you got the language or communication difficulties outlined but not some examples of developmental delays?

Group 5, page 4

Scale 5	Title	Clarity	ordering to be different as well [title]. For example, *problems with specific developmental delay in language or communication* because that implies language, communication and other specific developmental delays

Group 5, page 4

Scale 5	Description	Clarity	For example a child with Cerebral Palsy do you include them or exclude them, because it says

			exclude those with a disability, but in fact you include them as they've got motor problems. … Needs to be more detail about what specific developmental delays are included with examples of motor or coordination difficulties; examples are all language-based, so it makes it feel like that's all you're interested in …its confusing with Scale 6 about physical illness or disability …clarify motor or other specific developmental delays include…

Group 6, page 2

Scale 5	Rating	Wording	*Mildly severe* – how can you have those together? Suggestion: mildly significant

Group 6, page 2

Scale 5	Description	Clarity	Add: *fine and gross* motor…[first line]

Group 6, page 2

Scale 5	Description	Clarity	Confusing…*don't include infants with global learning disability*…but then up the top its say to put that in, irrespective of the cause or solution

Group 6, page 2

Scale 5	Description	Clarity	Less than 12 months – do not know if there is a global disability

Group 6, page 2

Scale 5	Description	Clarity	Bigger emphasis on language but it seems like you're looking at all areas of development - there's a lot more examples of speech and communication rather than any other kind of development

Group 6, page 3

Scale 5	Description	Clarity	Include social development …scale 10 is on social reciprocity

Group 6, page 3

Scale 5	Title	Clarity	Change title [no suggestion]

Group 6, page 3

Scale 5	Description	Clarity	Include examples of slow motor development

Group 6, page 3

Scale 5	Description	Clarity	Is this only about delay, or also include precocious development

Group 6, page 3

Scale 5	Description	Clarity	Include problems with slow processing – important to observe that

Group 6, page 3

Scale 5	Description	Clarity	Include limited use of words to express themselves; talk about expressive language …do you think the people using this would be all be familiar enough that you could use words like *expressive* and *receptive*? …include problems with understanding producing (communicating) and expressing language

Group 6, page 3

Scale 5	Wording	Clarity	Word *pragmatic* needs to be explained

Scale 6: Problems with physical illness or disability	
Include	Physical health problem or disability which limits or prevents movement, impairs sight or hearing or otherwise interferes with functioning. Problems in this area may be observed or based on reports from others. Movement disorder, side effects from medication. Congenital health problems. Physical health problems that are the result of psychological issues, deprivation, maltreatment or self-injurious behaviour. Include: side effects from medication, physical effects from drug/alcohol use, or physical complications of psychological disorders such as severe weight loss. Include physical complications or disability as consequence of self injury. Ratings will be influenced by consideration of impact of illness on everyday functioning.

Rating	Description
0	No incapacity as a result of physical health problems during the period rated.
1	Slight incapacity as the result of a health problem or disability during the period (e.g. cold, non-serious fall, teething). Parent voices concern about transient physical illness or physical symptoms but these are not considered serious by the parent or clinician.
2	The infant has some mild symptoms of physical illness or disability, that occasionally prevent engagement in usual activities, however the overall structure of their day is preserved and the ability to play is only mildly affected.
3	The infant is suffering moderate symptoms of physical illness or disability, resulting in some ongoing distress and loss of function. There is still some time each day, in which the infant/child is able to engage in usual activities, including play.
4	The infant is suffering severe symptoms of physical illness or disability that result in serious distress and/or loss of function. Normal everyday routines and activities, including play, are generally not possible because of the physical problem. Considerable input of effort and resources may be required to care for the infant, and support the parent.

Scale 6: Focus Group Responses

Group 2, page 5

Scale 6	Scale	General	Straightforward

Group2, page 5

Scale 6	Rating	Rating	Rating Scale is good; quite descriptive

Group 1, page 15

Scale 6	Description	Clarity	Does not mention self injury

Group 1, page 15

Scale 6	Wording	Clarity	The Scale references self injurious behaviour twice; double up

Group 2, page 5

Scale 6	Description	Clarity	Exclude: processing limiting physical activity, or coded in other item?

Group 3, page 4

Scale 6	Scale	General	No specific comments to make … Looks all right

Group 5, page 4

Scale 6	Description	Clarity	Similar to our comments on the previous one – motor developmental difficulties – where to put those

Group 5m page 4

Scale 6	Wording	Clarity	Change *drug and alcohol use* to *drug and alcohol exposure*

Group 5, page 4

Scale 6	Description	Clarity	Those physical illnesses such as epilepsy – are they captured here?

Group 5, page 4

Scale 6	Wording	Clarity	Language needs tidying up … *Movement disorder…* Then you've got *Side-effects from medication* which is completely different area yet again, *then congenital…*how its presented; – that's a lot …examples of movement disorders … *Physical health problems as a result of psychological issues*…I'm racking my brain – could be more clearly presented - make more explicit - physical health problems with no identified organic cause …its messy …random *physical complications of psychological disorders such as severe weight loss* - why isn't that in feeding behaviours? If it's in an infant, it's not like they have anorexia … Severe weight loss [should go in scale 4] nutritional problems … Examples need to be a lot clearer and presented in a logical fashion

Group 5, page 5

Scale 6	Rating	General	Descriptions of the ratings are quite good. I feel I could rate them if I was clear about what I had to include.

Group 5, page 5

Scale 6	Description	Clarity	The opposite has happened with *sensory processing* it's taken a lot of work to explicate to what that means

Group 6, page 3

Scale 6	Description	Clarity	*Physical effects from drug and alcohol use* – needs to be more clear …what about people who give their kids drugs? [Phenergan/Panadol; any antihistamine with sedative properties] …or all of our parents who dipped the dummy in the brandy or something! …would be dopey when you visit them as a side effect of medication

			…could mean from the parent or the child being given them

Group 6, page 3

Scale 6	Ratings	Clarity	[Rating 2]: *the infant has some mild symptoms of physical illness or disability, that occasionally prevent engagement* add *prevent or challenge engagement*

Group 6, page 3

Scale 6	Ratings	Clarity	[Rating 4]: *not possible* – change to "significantly or seriously impacted"

Group 6, page 3

Scale 6	Description	Clarity	Change *physical effects from drug/alcohol use* to distinguish between ingestion versus being given it

Scale 7: Problems associated with regulation and integration of sensory processing	
Include	Problems associated with processing, regulating and integrating information from sensory stimuli which interfere with the sensory regulation required for adaptive interaction with and exploration of the world. While problems with sensory organs are rated at scale 6, this scale is more concerned with the processing of otherwise apparently intact sensory organs. For example, the question is not whether their hearing is OK, but rather, is their processing of auditory information peculiar or problematic? Problems associated with sensory processing can reflect hypersensitivity (over-reactive therefore avoidant or fearful/cautious) and / or hyposensitivity (under reactive therefore seeking or impulsive) to one or more normal sensory stimuli. Sensory stimuli include vision, touch, hearing, taste, smell, and spatial awareness including the sensation of movement and awareness of body position in space. Problems associated with the regulation and integration of sensory processing are predictable and usually occur across multiple settings and within multiple relationships. Intensity, frequency, duration and location of problematic sensory stimuli may impact on the infant's presentation. Examples of the manifestation of sensory regulation difficulties include responsiveness to fabrics, movement, travel, focus on apparently irrelevant objects, avoidance of play, abnormal motor tone, floppiness in interactions with parents, frozen watchfulness. They may appear to have a preference for swaddling, or to seeking or avoiding certain fabrics. At more severe levels, the infant may be underweight and presents with failure to thrive due to discomfort and distress of feeding. Problems with sensory regulation may manifest as emotional dysregulation or emergent developmental delays.
Exclude	Problems with physical illness or disability rated at scale 6. Problems associated with specific developmental disabilities rated at scale 5. Problems with disruptive behaviour /under controlled emotional regulation rated at scale 1, while problems with anxiety and depression rated at scale 9. Activity levels, joint and/or sustained attention not appropriate to developmental age rated in scale 2.

Rating	Description
0	No sensory processing problems identified during the rating period.
1	Some concerns about sensory processing problems (over or under responding to normal sensory stimuli) however the impact on adaptive daily functioning and exploration of the world is minor.
2	One or more sensory processing problems identified and are impacting the infant. The infant/ and /or family may be showing signs of distress but maintaining appropriate developmental milestones. May become agitated, distressed, or disengaged when exposed to specific sensory stimuli. Definite and minor impact on functioning in daily tasks or in on maintaining interactions in primary care-giving relationships.
3	Sensory processing problems are impacting on the infant capacity to engage with the environment. May manifest as diminished exploration and play. Definite and moderate impact on daily functioning.
4	Severe and/or ongoing difficulties related to sensory processing problems directly impacting the infant social, emotional and physical wellbeing. Definite and severe impact.

Scale 7: Focus Group Responses

Group 2, page 5

Scale 7	Description	Clarity	Wordy, particularly with examples at the end

Group2, page 5

Scale 7	Description	Clarity	Our team thought there was great detail about the sensory components, but perhaps *abnormal motor tone, floppiness and frozen watchfulness* are such broad issues and may not reflect sensory processing abnormalities, it may be better not to put that in.

Group 2, page 5

Scale 7	Description	Clarity	Err on the side of being more prescriptive …there or in glossary

Group 3, page 4

Scale 7	Description	Clarity	Take out “abnormal motor tone”, would go into an example of a motor difficulty in the other scale, and isn’t specifically related to sensory

Group 3, page 4

Scale 7	Description	Clarity	Presumably you’re excluding feeding sensitivities from here [suggested to say that in exclusions] would come under feeding

Group 3, page 4

Scale 7	Description	Clarity	Sensory processing could go to sensory regulation

Group 4, page 5

Scale 7	Education and training or instruction	How to rate	This is the difficult one

Group 4, page 5

Scale 7	Scale	How to rate	If it was a specific issue, I am thinking OT-wise, that could go under specific developmental delay; OT’s

			look at sensory integration issues – over-reactive or under-reactive …the two OT's who work with us would be marvellous to review this tool

Group 4, page 5

Scale 7	Scale	description	Include sensory auditory processing issues

Group 5, page 5

Scale 7	Description	Clarity	Good that there is a lot of information included for people who are not familiar with this field, it needs to be spelt out

Group 5, page 5

Scale 7	Wording	Clarity	*Problems associated with the regulation and integration of sensory processing* [delete:] predictable

Group 5, page 5

Scale 7	Description	Clarity	Did not like *for example, the question is not whether their hearing is OK, but rather, is their processing of auditory information peculiar or problematic?* Use visual, tactile, sensory – need examples; how its experienced. Examples of good *manifestations of…* That kind of everyday language would be more useful.

Group 5, page 5

Scale 7	Wording	Language	Do not use *OK* - sounds too familiar. Rest is technical, this is too familiar.

Group 5, page 5

Scale 7	Rating	Clarity	Descriptions of each rating were good

Group 6, page 3

Scale 7	Education and training or instruction	How to rate	The difficult thing with diagnosing a sensory problem is that we don't know whether it's a biological/emotional problem or due to the child being emotionally dysregulated

			…not aetiological – its what's happening with the infant …not talking about diagnosing, but what factor it has on the child's health … May indicate emotional/biological difficulties …biological rated at Scale 5 rather than here? … Clarify that it may be associated with underlying difficulty processing their emotions …could be indigenous to the child or their environment; rate both here – as the second sentence in paragraph 1 of description

Scale 8: Problems associated with sleep	
Include	Sleep disturbance is common for infants. Include difficulties in both settling and maintaining sleep irrespective of where the locus of the difficulty is thought to be (infant, parent, living arrangements). Include excessive sleep (e.g. which interferes with opportunities for skills or social development), insufficient sleep (e.g. periods of awakenings or reduced sleep time) and disturbed sleeping (e.g. sleep talking, sleep walking, night terrors, or any other disturbance during sleep when the infant does not seem to respond to the parents). Include snoring or loud mouth breathing with breath holding or gasping.

Rating	Description
0	No problems of this kind during the period rated.
1	Minor problems, typically within expected developmental norms, infrequent and where the family appear to have some approaches that successfully address the problem.
2	Mild problems which are intermittent and the family appear to have some success in addressing the problem.
3	Moderately severe sleep problems. The infant‘s sleeping pattern is a cause for distress in the parents and family. The sleep disturbance is present most of the time, and may be significantly out of keeping with age expectations.
4	Severe sleep problems. The sleeping pattern is a cause for great distress in the parents and family and may be significantly out of keeping with age norms. The sleep disturbance is present nearly all the time and significantly interferes with development.

Scale 8: Focus Group Responses

Group 1, page 15

Scale 8	Wording	Clarity	[Rating 3] *The infant's sleeping pattern is a cause for distress in the parents and family.* Depending on parent - Parent can't sleep, or can't cope. Seems unfair to rate the child according to the parent. It may only matter that the infant has a sleep issue; it may not matter what the cause of the issue is. … Parent may have a low tolerance. Is the scale is supposed to pick that up or not?

Group 1, page 15

Scale 8	Rating	Measure	There is a big jump between 2 and 3 on this scale and not much of a gap between 3 and 4.

Group 1, page 15

Scale 8	Description	Clarity	Give more examples in the description like insufficient sleep for younger infants. Most of the examples in brackets are for older babies. For example "difficulty resettling" for junior babies.

Group 2, page 5

Scale 8	Description	Clarity	Good - clarified that it could be with the parent or the environment

Group 2, page 5

Scale 8	Description	Clarity	Not sure what that first sentence is supposed to imply – normalising.

Group 2, page 5

Scale 8	Description	Clarity	*Disturbance* – change to *difficulties*

Group 2, page 5

Scale 8	Description	Clarity	[Rating 3] *The infant's sleeping pattern is a cause for distress in the parents and family.* Will depend on age; if older and still highly disturbed, its problematic. Put more explanation around that.

			[Suggestion:] “Sleep patterns can vary significantly with age.”

Group 3, page 4

Scale 8	Education and training or instruction	Rating	With scale, but so subjective; you could get such a range of ratings – for example – could get a kid of six times a night and it’s fine, or a kid up three times a night and it’s a disaster, and that would be parent rated and also clinician rated. I think there’s a lot of clinician prejudice in how they might rate this scale in particular.

Group 4, page 5

Scale 8	Scale	Description	Snoring would be physical?

Group 4, page 5

Scale 8	Scale	General	Scale is good

Group 5, page 5

Scale 8	Wording	Description	Change *disturbance* - I see “problem”; threw me off

Group 5, page 5

Scale 8	Description	Clarity	Co-sleeping should be included but not put in a way that it pathologises it - open to interpretation … Could be cultural – may not matter – still a problem … Put in HoNOSCA and all scales

Group 5, page 5

Scale 8	Rating	Clarity	[Rating 4] include *age “*and cultural” *norms*

Group 5, page 5

Scale 8	Description	Clarity	Included night terrors but that is very normal… Maybe none of them... Parasomnias… need to be included [ie. Exclude them, esp night terrors … Capacity to settle independently should be included

Group 6, page 4

Scale 8	Description	Clarity	Weighted very much on parental ability to cope with it – does not relate necessarily to the sleep problem. You may have a very relaxed parent who couldn't care less, or you might have an uptight parent who can't cope with one waking a night. … Use some measure other than whether it interferes with the family

Group 6, page 4

Scale 8	Rating	Clarity	Does it need to be that its a rating of 4 before it impacts on the child's development?

Group 6, page 4

Scale 8	Rating	Clarity	*The sleeping pattern is a cause for great distress in the parents and* "interferes with the child's functioning" or *something*.

Scale 9: Problems with emotional and related symptoms	
Include	Symptoms of depression, anxiety and phobias. Problems with negative or inhibited affect in the infant suggestive of low mood, anxiety, fear or emotional withdrawal. May include fears, anger, anxiety or emotional withdrawal from parents and others. Include incongruent lack of emotional expression. May be expressed with changes in curiosity, clinging, masking face, incongruent emotional expression, startle responses, crying, anger, hypnotic gaze, withdrawal, and blank expression, exaggerated positive or negative emotional responses. May include excessive stillness, frozen watchfulness, quiet rage and restrictions in affect range. Apparent increased tolerance for aversive adult behaviour should be considered. Include emotional responsiveness to challenges in the environment and requirement for comfort. Include age or developmentally inappropriate lack of wariness, or avoidance of parents.
Exclude	Physical sequelae of psychological disorders or medication – rated at scale 6. Disruptive behaviours resulting from emotional distress – rated at scale 1. The emotion associated with the disruptive behaviour is rated here at scale 9.

Rating	Description
0	No problems in the rated period.
1	Transient and/or minor mood, anxiety and emotional symptoms or changes.
2	Mild but definite emotional symptoms are present but not preoccupying.
3	Moderately severe emotional symptoms are present.
4	Severe emotional symptoms which intrude into all activities and are nearly always uncontrollable.

Scale 9: Focus Group Responses

Group 1, page 15

Scale 9	Description	Clarity	Not focusing on junior [under one year] babies that much. … Harder for your average clinician to see…thinking of CYMHS using this tool. You can relate 2-year old behaviour to something but to do that with a week-old baby or month-old baby is a very difficult picture.

Group 2, page 5

Scale 9	Scale	General	Trying to encompass an extraordinarily broad range of things – half of the DSM in one scale

Group 2, page 5

Scale 9	Description	Clarity	I think clinicians will be able to rate that. You don't need to be overly inclusive. Its something that we're pretty familiar with.

Group 3, page 5

Scale 9	Description	Clarity	May include changes in curiosity, emotional responses, excessive stillness and restrictions in affect range, increased tolerance for aversive adult behaviour [make the inclusions dot points]

Group 3, page 9

Scale 9	Scale	General	I actually like this one. I think its pretty good … Has different developmentally appropriate examples – I think that's good … Encapsulated well…Delineated examples

Group 4, page 1

Index of scales	Scales	Structure	Eye gaze should be in Scale 9

Group 4, page 5

Scale 9	Description	Clarity	[scenario described where a child was removed by protective services]…and everybody said "Isn't he good" but his reaction was *extreme*…and it was

			severe…overly controlled…internalisings…child is at enormous risk – [scale] does not describe the child

Group 4, page 5

Scale 9	Description	Clarity	This scale is a mish-mash

Group 4, page 5

Scale 9	Scale	How to rate	Shut down to me means a different thing, than a good child doing the appropriate behaviour …I don't feel this tool captures this in any way

Group 5, page 6

Scale 9	Scales	Where to rate	*Settling, demanding, wining, irritability, excessive crying* – that whole paragraph would probably be better in scale 9 because they're not disruptive behaviours; they're signs of a child distressed, whereas biting, hitting, being defiant – they're behavioural

Group 5, page 6

Scale 9	Scales	Where to rate	Haven't got non-organic symptoms here. Eg. if a child is still soiling inappropriately, not due to physical or mental impairment, could be a sign of rage - that would be in the disruptive behaviour one

Group 5, page 6

Scale 9	Description	Clarity	Confusing – exclusions – *physical sequelae of psychological disorders or medication* you are going to have physical sequelae in young infants; that's how they show distress … Medication maybe, but the other bit does not make any sense … Trying to think what that might be

Group 6, page 4

Scale 9	Description	Clarity	Add *trauma* after *fear* in first paragraph

Group 6, page 4

Scale 9	Description	Clarity	Add *trauma* after *fear* in first paragraph

Group 6, page 4

Scale 9	Description	Clarity	[NB: – detailed, specific discussion to get these words right:] Change to: *"Include age or developmentally inappropriate lack of appropriate comfort-seeking or safety-seeking, lack of wariness or avoidance of adults/others, and disinhibited behaviour"*

Group 6, page 4

Scale 9	Description	Clarity	Include age-inappropriate sexualised behaviours with self or others

Group 6, page 4

Scale 9	Description	Clarity	After *curiosity* "decrease in exploration and play"

Group 6, page 4

Scale 9	Rating	Clarity	Jumps seem excessive – 2 to 3, then 3 to 4 don't seem even enough

Scale 10: Problems with social reciprocity	
Include	This scale addresses the capacity of the infant to engage in, and engage others in, age and developmentally appropriate interactions. There may be problems with seeking, engaging and enjoying interactions with familiar adults and children, including development of the social smile at 6 weeks. Responses to social engagement or social intrusion from others are not responded to appropriately e.g. ambiguous half smiles. Problems may manifest in reciprocity in communication, play, and games. Reciprocity may be expressed both pre verbally and verbally, as well as behaviourally. Problems may manifest as indiscriminate and overfamiliar social interactions as well as withdrawn and disengaged social interactions. The scale includes problems with emotional attunement and misalignment between the infant's needs and the parents' or carers' responses. Problems rated in this scale may include the infant's capacity to manage appropriate eye contact e.g. the infant may not gaze at the parent's face or at interesting object when shown. Problems may include avoidant gaze; no eye contact (but no active avoidance either); brief glances without sustained looking (difficulty gaining and sustaining eye contact); and unfocused eyes. Problems with vocalisations relating to reciprocity of interactions, such as turn taking, engagement attempts, and vocal mirroring may also be relevant indicators of social reciprocity issues.
Exclude	Difficulties with vocalisation separate to the social reciprocity function are rated at Scale 5.

Rating	**Description**
0	No problems during reporting period.
1	Transient or mild problems in the infant's developing capacity to engage in social relationships.
2	Mild but definite problems in at least one area.
3	Moderate problems in more than one area.
4	Severe issues with social reciprocity. Problems likely to occur in most areas and intrude across most interactions.

Scale 10: Focus Group Responses

Group 2, page 5

Scale 10	Scale	General	I think it's a really good Scale; I'm glad it's included.

Group 1, page 16

Scale 10	Description	How to rate	Need more examples around older babies [of social reciprocity]

Group 2, page 5

Scale 10	Rating	Clarity	There are two jumps that happen within that scale: you go from [Rating 2] *Mild but definite problems in at least one area* to [Rating 3] *Moderate problems in more than one area.* Its like its sort of fused - you go from mild, moderate, severe, and how many areas, sort of fused into one. … Or just drop the whole thing. Have mild problems, moderate problems, severe problems.

Group 3, page 5

Scale 10	Description	Clarity	How is the infant engaging in early relationships or is this kind of an early sign for autism kind of scale? [Preference – not to be an autism scale]

Group 3, page 5

Scale 10	Description	Clarity	Descriptors could be a little more specific. Look at it as more of a temporal concept, rather than one area: "The infant chose to engage in reciprocal social interaction at some time with some people", rather than in one area, or "the infant showed no capacity to engage interactions with anyone". A bit more specificity would help the rater.

Group 4, page 5

Scale 10	Scale	General	Like it [Scale] ..that one's all right

Group 5, page 6

Scale 10	Title	Clarity	Add: *and interpersonal sensitivity*

Group 5, page 6

Scale 10	Wording	Clarity	*includes problems with* "parents" *emotional attunement* "and their responses to infant's needs"

Group 5, page 6

Scale 10	Description	Clarity	Including more paragraph breaks – helpful when you're reading things [referring to the last paragraph of inclusion box on scale 10]

Group 6, page 4

Scale 10	Description	Clarity	*Eye contact* also use social referencing, to check what someone is thinking or feeling Change to *eye contact or to socially reference others*

Scale 11: Problems with age appropriate self-care and environmental exploration	
Include	This scale addresses age-appropriate self-care and exploration of the environment. Self-care is likely to be a more prominent consideration with older children. Self-care is likely to include age appropriate levels of assistance with bathing, feeding, dressing, playing etc. Problems with contributing to self-care and environmental exploration often exist due to restrictions in the infants' environment. It is critical to rate the problem and good practice to not assume that the rating indicates that the infant is the sole source of that issue. Include problems with basic activities of self-care such as washing, dressing, toileting. Exploration may include visual, tactile, verbal as well as physical exploration (under or over exploration). Include problems with complex skills such as play, autonomous activities or separating from parents, taking into account the norm for the infant's age and developmental stage. Difficulties may be indicated by regression to an earlier stage of development. Problems with separating from parents when the infant is attending structured socialisation settings (e.g. day care, pre-school) may be rated here although the resulting impact on attendance at socialisation settings should be rated at scale 13. Include poor levels of functioning arising from apparent lack of motivation, mood, environmental restriction or any other issue whether it is considered to arise from the infant or the parents or the environment.
Exclude	Do not include feeding problems rated at scale 4. Do not include sleeping problems rated at scale 8. . Do not include lack of opportunities for exercising intact abilities and skills, as might occur in an over-restrictive family rated at scale 12. Do not include the outcome of limited environmental exploration on structured socialisation settings rated at scale 13.

Rating	Description
0	No problems during reporting period.
1	Minor problems with self-care or exploration of the environment.
2	Mild problems with self-care or exploration of the environment.
3	Moderate level of problems with self-care or exploration of the environment.
4	Severe level of problems with self-care or exploration of the environment that is likely to be intruding across settings, activities and persons.

Scale 11: Focus Group Responses

Group 1, page 16

Scale 11	Description	Clarity	*Self-care* and *exploration* are such different topics. … Self-care might be about the older baby. Exploration may be about the younger baby; Might be looking at the same thing, with the age differential? So maybe that needs to be clarified in the statement, if that's is correct.

Group 1, page 16

Scale 11	Scale	Useful	Environmental exploration is a good topic to have as a scale; helpful.

Group 3, page 5

Scale 11	Description	Clarity	Tricky one… Its adaption from all the children and adults… It's a clunky concept in my mind. … There's a few things – it's toileting [regulatory difficulties], adaptive behaviour [adaptive self-care involves feeding], which we've covered off elsewhere. Environmental exploration – you can get the trauma pattern… or the aloof disconnection… So clinically if you think about the clusters of symptoms you would see in this category, it draws on a lot of different possibilities ... The other thing is, from preschool, self-care is such a relational concept, this struggles to capture the relational element of it; can this person seek help from someone to get dressed? Can they signal their need to go to the toilet?… That's where it grates. It probably lacks that relational element, and are we measuring something else if we include that?… My worry is… If you get a kid who's really avoidance in their attachment and does everything fully independently, then that might rate really highly on this scale and yet its quite maladaptive because they're not able to use relationships to get their needs met. … Over reliance or under reliance on caregivers for support

Group 3, page 5

Scale 11	Description	Clarity	Have some early examples for little ones, particularly of exploration… In that earlier age range, you're looking for exploration, not self-care.

Group 3, page 5

Scale 11	Title	Clarity	*Can they use caregivers to meet their needs?* That would be a better title

			[agreement heard]

Group 5, page 6

Scale 11	Scale	General	Routine … Fine

Group 5, page 6

Scale 11	Description	clarity	would not say *it is critical to rate the problem and good practice to not assume* I would just say: it is vital to pay attention to the fact that infants… Something like: Infant's difficulties may be related to the parent's capacity …later confused… *Do not include lack of opportunities for exercising intact abilities and skills as might occur in an over-restrictive family rated at scale 12* - how that overlaps with what we've just been talking about – helicopter – need the caregiver to scaffold it for you in order to be able to do it

Group 6, page 5

Scale 11	Title	Clarity	Add: *in an infant*

Group 6, page 5

Scale 11	Description	Clarity	Difficult to assess in an infant. Eg 2-month old – I just rated as a zero

Group 6, page 5

Scale 11	Description	Clarity	Add: Environmental expression in an infant may be expressed as visual curiosity, body movements to show excitement

Group 6, page 5

Scale 11	Description	Clarity	Add: restrictions or conditions of environment in that second paragraph eg. kids who have been in detention

Scale 12: Problems with family life and relationships	
Include	This scale addresses problems in family life that are thought to impact on the infant. If the parents are separated, consider the relationship with each parent and the separated parents' ability to co-parent where appropriate. Include relationships with significant others should be included – grandparents, siblings, extended family members, child care providers. Include instances of neglect including physical (e.g. lack of sufficient access to appropriate food, shelter & clothing) and emotional (e.g. lack of warmth, comfort and age appropriate regulation of the infant's affect). The availability of access to caring attentive and empathic adults, and the ability to keep the infant in mind, should be considered. Include parent or family irritability with the infant. Difficulties in managing powerful emotions or any consequent harmful behaviour by those in the infant's immediate environment should be considered. Include instances of physical or verbal hostility or abuse towards the infant, as well as family hostility or conflict which impacts on the infant. Consider capacity for significant others to contain powerful negative emotions towards the infant. Issues such as parental or sibling mental health, substance use and personality problems should be included if they have an effect on the infant.
Exclude	Do not include disruptive behaviour by infant, rated at scale 1. Do not include problems with social reciprocity rated at scale 10.

Rating	Description
0	No problems during the period rated.
1	Slight or transient minor problems. Some concerns about family relationships are evident but these are part of a history where reasonable parental reflective capacity is typically apparent.
2	Mild but definite problem that have some impact on the infant's development.
3	Moderate problems in the relationship between the infant and primary caregivers or family that are affecting the infant's regulatory and exploratory capacity.
4	Serious problems in family relationships with severe impact on the infant.

Scale 12: Focus Group Responses

Group 1, page 16

Scale 12	Rating	Clarity	[Rating 1] *Parental reflective capacity* needs to be well defined. Do untrained people know what this is? Spell out more.

Group 1, page 16

Scale 12	Rating	Clarity	Distil it further between rating 2 and rating 4. In rating 2 there is a problem in the relationship for the child, but there is another significant person like at pre-school or grandparent, whereas in rating 4 - there is a serious problem in the relationship plus there's no one else … No functional relationship

Group 2, page 6

Scale 12	Description	Clarity	No use of the word *trauma* in there.

Group 3, page 6

Scale 12	Scale	General	I like this scale

Group 3, page 6

Scale 12	Description	Clarity	I wouldn't mind seeing *parental reflective capacity* which is mentioned [Rating 1]… Either put in the more severe ratings, or remove from here and put in the descriptor; to make it clear that its practical physical resources as well is parental reflective capacity that are being rated on this scale

Group 4, page 6

Scale 12	General	Scale	Fine

Group 4, page 6

Scale 12	Rating	How to rate	Score 4 [patient scenario described] – mothers' view of the child rather than her behaviour towards him – this may not capture the child [suggested change to:]

			"*Likely* to have severe impact" [where there isn't a severe impact] – *Would be* likely, because we can see that - should be coded as a problem – risk, potential impact – if we change the wording, it could improve it

Group 5, page 6

Scale 12	Scale	General	These issues are done better than the HoNOSCA eg irritability. Scapegoating – nowhere neat to put that in the HoNOSCA

Group 6, page 5

Scale 12	Description	Clarity	First paragraph – highlight also siblings

Group 6, page 5

Scale 12	Ratings	Clarity	Ratings – OK

Scale 13: Problems with attending care, education and socialisation settings	
Include	This scale addresses attendance at the prime structured socialisation setting outside of the family. Include attendance at any type of regular socialisation and care activity at the time of rating e.g. regular care with extended family or formal early childhood education (sometimes called kindergarten or pre-school). Include activities irrespective of location e.g. regular play group sessions at infant's home. Include refusal of, or withdrawal from early childhood education, childcare, play group or similar regular socialisation activity, irrespective of cause. Include limited or minimal opportunities to attend socialisation activities appropriate to the infant's age. Include consideration of additional supports such as reassurance, transitional objects, required to settle the infant in the setting. If early childhood education, childcare etc. is in holiday break, rate the last two weeks of the previous open period. Note: Infants and young children will communicate their reluctance and distress at attending these settings through a range of symptoms. These may include problems in feeding, toileting, eating, playing, communicating and sleeping both at the settings and around the transition time. These symptoms in themselves are likely to be rated at different HoNOSI scales and are not the sole source of rating at this scale. However, it is acknowledged that the reluctance to attend may be conveyed to the clinician through these symptoms. The actual attendance problems are rated at this scale.
Exclude	All behaviours and emotional expressions of problems associated with attendance or separation are rated at their respective scales (e.g. Disruptive at scale 1, Feeding at scale 4, Emotional at scale 9). Absences due to illness of infant or parents requiring them to be absent from the setting. This typically includes medical conditions, such as fevers, contagious illnesses or infections which would be rated at scale 6.

Rating	Description
0	No problems during the period rated. Infant displays age appropriate behaviour on separation from their parents and settles readily when comforted in the environment.
1	Minor problem with attending and may display reluctance for brief periods. Responds with small amount of support additional to that typically required at this age.
2	Mild but definite problem with some sessions missed or refusal to participate in activities when attending.

3	Moderate problem with several days missed during rating period due to infant's reluctance to attend.
4	Severe problem with infant absent for most of the days or sessions during rating period.

Scale 13: Focus Group Responses

Group 1, page 16

Scale 13	Rating	Clarity	If the parent is not bringing the child to playgroup or any other social setting it would be rated as a 0, but it should not be a 0. … They talk about the exclusion in scale 11, but maybe it needs to be included in 12 and 13. … Or maybe the parent not taking the child needs to be included in rating 1,2,3, or 4 [agreement] *Include refusal, include limited or minimal opportunity…* Rating is about the infant, not the mother. Eg. if the infant is not getting the opportunity to socialise [daycare, cousins, friends]

Group 2, page 6

Scale 13	Scale	General	Straightforward

Group 3, page 6

Scale 13	Ratings	Clarity	The descriptions down here[referring to ratings] need to be very vague because it may be that the child is not going to anything and that is absolutely appropriate, say you'd have zero. It may be that the child needs five days a week childcare and is just not getting there and that may be a four; it's not about attendance it's about appropriateness that the child is getting. … Don't make it the absence of, make it the inappropriateness of [attendance]

Group 3, page 6

Scale 13	Title	Clarity	Problems with appropriate engagement in non—parent care

Group 5, page 6

Scale 13	Scales	Where to rate	Some of that stuff that included exploration may be doesn't need to, if it's done here

Group 5, page 6

Scale 13	Scales	Clarity	Big section may not be warranted. If not going to a childcare centre … its about can they cope in a social situation apart from the one of their primary care giver … Focus on educational care settings feels like its done to stick with item 13 on the HoNOSCA, rather than what's developmentally appropriate

Group 6, page 5

Scale 13	Description	Clarity	Didn't really apply in an infant still at home. Add: Do not rate very young infants – don't want to give them a low mark because they're very young and do not go to play group …so that would be a 9

Group 6, page 5

Scale 13	Description	Clarity	…include not just opportunities, but their ability to use opportunities

Group 6, page 5

Scale 13	Rating	Clarity	[Rating] Opportunity versus reluctance - *refusal* due to anxiety?

Group 6, page 5

Scale 13	Description	Clarity	Change to *Include limited or minimal opportunities* and capacity *to attend* and participate in *socialisation activities*

Group 6, page 5

Scale 13	Title	Clarity	Problems with *participation* and attendance

Group 6, page 5

Scale 13	Rating	Clarity	Make rating consistent: minor, mild, moderate

Scale 14: Problems with knowledge or understanding about the nature of the infant's difficulties	
Include	Include lack of useful information or understanding available to the parents, caregivers or referrers or support system about the nature of the difficulties. Include problems with capacity or knowledge to understand the infant's difficulties. Include limited or incorrect understanding of the infant's developmental stage and needs. Include misunderstanding, minimising, elaborating or exaggerating the difficulties, impact or distress as well as inaccurate attribution of the infant's difficulties. Include lack of explanation about the difficulty/diagnosis, the cause of the problem or understanding of the prognosis or the impact on the infant. Rating a problem here does not preclude that the service system may revise their understanding of the infant's difficulties. In many ways, problems rated here may indicate a lack of congruence between the parent's and other key figure's views about the nature of the difficulties and the assessing or treating system's views.

Rating	Description
0	No problems during the period rated. Parents demonstrate a good level of understanding about the difficulties.
1	Minor problems only. For example, parents/ essentially understand infant's difficulties but with occasional misunderstandings such as sometimes downplaying, or exaggerating the infant's difficulty or distress.
2	Mild but definite problem in understanding infant's difficulties.
3	Moderately severe problems. Parents have very little or very poor knowledge about the nature of their infants' problems.
4	Very severe problem. For example, parents have no understanding about the nature of their infant's problems.

Scale 14: Focus Group Responses

Group 1, page 16

Scale 14	Rating	Clarity	Normalising – first-time mum with newborn infant won't know, so we don't want to give a zero [distinction between that and major lack of knowledge]

Group 1, page 16

Scale 14	Rating	Clarity	[Rating 0]: *parents demonstrate a good level of understanding about the difficulties* suggest to add: "or where to go to get information".

Group 1, page 16

Scale 14	Description	Clarity	With this scale I find it hard to separate those parents who kind of know that there is a problem, but don't know where to go [from rating that at Scale 15]… Some mums may have difficulty recognising there is a problem… I think exaggerating difficulties is important too.

Group 2, page 6

Scale 14	Scale	General	I think it just forces people to ask a question that some of us routinely ask anyway "tell me what you think's going on"

Group 3, page 6

Scale 14	Wording	Clarity	Change to: *problems with capacity to understand or knowledge to appreciate the infant's difficulties*

Group 3, page 6

Scale 14	Wording	Clarity	Put *consider* instead of *include* the whole way through

Group 4, page 7

Scale 14	Wording	Clarity	Lack of useful information is a bit of a furphy [suggested its not needed] -how would you know that? Not having been given the information is

			different from not having the capacity to take it on and support the infant appropriately

Group 4, page 7

Scale 14	Scale	General	Scale is all right

Group 5, page 7

Scale 14	Scales	Clarity	I never understand the difference between 14, 15

Group 5, page 7

Scale 14	Wording	Clarity	*Include problems with capacity or knowledge to understand the infant's difficulties* add "and needs"

Group 6, page 5

Scale 14	Scale	General	Great

Group 6, page 5

Scale 14	Rating	Clarity	Standardise the rating system

Scale 15: Problems with lack of information, understanding about services, or managing the infant's difficulties	
Include	Include lack of useful information available to the parents, caregivers, or referrers, or a lack of understanding regarding services or management of the difficulties. Include lack of useful information or lack of understanding about the most appropriate social, therapeutic, developmental, care, and educational supports and interventions for the infant's presentation. Include here supports and interventions that the parents could be providing directly to their infant. Include parental willingness to access/utilise services and interventions to support the infant. The consistency with which parent's understand or use appropriate management strategies, and the extent to which supports are required to help the parent's use optimal approaches may be considered here. Include information not acted upon such that appropriate and feasible interventions are not implemented adequately. Include family's unwillingness or social isolation if it restricts engagement with appropriate services. Child protection authorities may be involved if the match between lack of engagement with the provided information and the infant's difficulties are severe enough. Note: The lack of useful information may stem from a failure of the environment to provide such information or from the parent's misunderstanding or unwillingness to act on available information. Include parent's reflective function regarding their role in supporting the infant and their ability to engage with and utilise services available. Rating a problem here does not preclude that the service system may revise their understanding of the optimal approach to managing the infant's difficulties. In many ways, problems rated here may indicate a lack of congruence between the parent's and other key figure's views about the management of the infant's difficulties and the assessing or

Rating	Description
0	No problems during the period rated.
1	Minor problems only. For example, parents have a good enough understanding of how best they and other resources can help their infant, or they are actively seeking appropriate information, support or access to services.
2	Mild but definite problem in understanding the appropriate services, approaches, resources and supports for the infant's difficulties.
3	Moderately severe problems.
4	Very severe problem.

Scale 15: Focus Group Responses

Group 1, page 16

Scale 15	Rating	Clarity	Rating 1 description matches a rating for 0.

Group 1, page 16

Scale 15	Scale	Rating	Needs to include if they have the knowledge, and willingness to go, but do not follow through or can't get there.

Group 1, page 17

Scale 15	Description	Rating	Needs to include that cohort of parents who know what they need to do, they know they need help, but they are so scared about what that might mean, they then don't. May have had a negative experience in the past… If I ring mental health, that means FACS.

Group 1, page 17

Scale 15	Description	Rating	What about parents with religious beliefs – may impede on assessing services … There are lots of other reasons people would not access support; it's not just about lack of information out there.

Group 4, page 7

Scale 15	Scales	General	Another one that has been brought down from the other scales and its from a disability model in the HoNOS. What might be more relevant is that parents aren't accessing the supports and services that would be helpful for their infants, or its not available

Group 5, page 6

Scale 15	Description	Clarity	*Lack of understanding of the most appropriate intervention* – should be explicated that it's not

			parents' lack of understanding, its systems lack of understanding

Group 5, page 7

Scale 15	Rating	Clarity	Give more examples of *moderately severe* and *very severe*, like in scale 14 to explain how they are different

Group 6, page 5

Scale 15	Scale	General	Outstanding – no need to be modified

Appendix 4: Additional Responses from Non-Focus-Group Settings - obtained via Teleconference, Videoconference or Individual Clinician Interview

This section shows the responses from the non face-to-face participants who added additional information. Participants comprised of 1 group of 3 people via teleconference, 1 person via telephone interview, 2 people via telephone interview that stopped at Scale 6 due to time restriction, 1 person via Skype interview and 1 person via face-to-face interview, for a total of 8 people. They were a combination of one or more Social Workers, Psychiatrists and Occupational Therapists. The findings, labelled as “Additional Responses” along with written responses received after the face-to-face focus groups were conducted, are outlined below under a copy of each relevant section of the HoNOSI.

The HoNOSI

Scale 1	Problems with disruptive behaviour /under controlled emotional regulation
Scale 2	Problems with activity levels, joint and/or sustained attention
Scale 3	Non accidental self-injury
Scale 4	Problems with feeding behaviour
Scale 5	Problems with language or communication skills or specific developmental delays
Scale 6	Problems with physical illness or disability
Scale 7	Problems associated with regulation and integration of sensory processing
Scale 8	Problems associated with sleep
Scale 9	Problems with emotional and related symptoms
Scale 10	Problems with social reciprocity
Scale 11	Problems with contributing to self-care and environmental exploration
Scale 12	Problems with family life and relationships
Scale 13	Problems with attending care, education and socialisation settings
Scale 14	Problems with parent's knowledge or understanding about the nature of the infant's difficulties
Scale 15	Problems with lack of information or understanding about services or management of the infant's difficulties

Additional Responses: Index of Scales

- Might not start at 1.
- Do not start with deficits. Could become cynically cold tool.
- Would not combine emotional regulation with disruptive behavior – tone is haughty “disrupting the peace”. May have difficulty managing feelings; not always disruptive behavior. Could underlie, but not necessarily
- Scale 1 overlap with Scale 9?
- Scale 11 – confused – odd combination: *Self care* and *environmental exploration*
- Scale 14 and 15 – overlaps? – *nature of the infant’s difficulties* related to *information or understanding about services or management.* Parents with high functioning but no knowledge tend to do better.
- Parents low functioning – “No problem there”, She has only 1 hour. She sees bit of a hint because rating is for previous 2 weeks – relying on parent report. My observation – where does that leave you as an observer?
- Other resources may be biased – put something up front? Eg Interview of parent + 1 hour observation of child.
- Covers large breadth of areas
- Modified from HoNOSCA

Key principles for rating the HoNOSI

- The rating period is the previous two weeks.
- Rate each scale in order from 1 to 15.
- Use all available information in making your rating. That is, HoNOSI ratings should reflect your judgement based on all sources of information available to you.
- Do not include information already rated in an earlier scale.
- Rate the most severe occurrence of the problem in the rating period, not the inferred cause.
- Clinically significant symptoms should be rated at a 2 or above.
- Ratings are informed by familiarity with, and a good understanding of, infant and child development.
- The instrument is focussed on the infant but any rating may also reflect the relationship with the parent(s). Include the infant's temperament, parent's responses, interactions and level of distress regarding aspects of the infant's behaviour and presentation. As with HoNOSCA, the presence of a clinically significant rating (scoring 2, 3 or 4) does not imply that the source of the problem, and/or the locus of intervention, is necessarily exclusively with the infant.
- When establishing a rating point, it can be useful to consider the underlying construct of a problem as a continuum.
- Each item is rated on a five-point scale of severity (0 to 4) as follows:

Score	**Rating**	***Suggested implications***
0	No problem.	
1	Minor problem requiring no formal action.	Sub-clinical problem, evidence of some behavioural disturbance or distress, unlikely to be monitored or included in care plan

2	Mild problem.	Mild problem, clinically significant issue evidence of distress and/or behaviour disturbance. Likely to be monitored or included in care plan
3	Problem of moderate severity.	Moderate problem, clinically significant issue evidence of greater distress and/or, behavioural impact. Definitely monitored and included in care plan
4	Severe to very severe problem.	Severe clinical problem, distress and/or behavioural disturbance dominant aspect of presentation, Greater frequency and/or intensity of clinical activity as evidenced in care plan
9	Not known or not applicable	

- Higher ratings can be expected to accompany more severe, more frequent and more widespread presentations.

- As far as possible, the use of rating point 9 should be avoided, because missing data makes scores less comparable over time or between settings.

- Specific information on how to rate each point on each item is provided in the Glossary.

Additional Responses: Key Principles

- It is important to infer cause - emphasise stronger perhaps
- Do need a good understanding of developmental milestones – hard to make a severity rating
- Even with professional disciplines, not agree moderate/severe – paediatrician – OK.
- Like the heavy rating on important domains – scale tends to focus on motor/physical
- To cover all 15 scales - really big, really demanding, could not cover it in the 1 hour that I have. Cluster? Include open-ended questions to help you to rate.
- "Explanatory notes are great"
- Jumped right into it – did not need the key principles.
- *Do not include information already rated in an earlier scale.* – confuses. First scale – emotional regulation infant injuring themselves does not cause harm. I would get confused is can't use it in another scale. Harming – might be redundancy in some of those questions.
- In ratings: *requiring no formal action.* – why is this here? None of these state requiring action. Might require action based on other scales once you have accumulated multiple, you may need to act. This may bias people from choosing number 1 [rating].
- Rating 1 has "no" on the left and "unlikely " on the right - not consistent

- Rating 9 – Not known or not applicable – why is that there? Does HoNOSCA use N/A – whether don't know or couldn't apply
- Guide for weighting rating according to observer/parent/other sources – can be biased regarding parents
- Developmental Milestones – is it like Ages/Stages [questionnaire]?
- 28 week prem baby would mean two months difference in development – in this age group every month counts
- as it stands it's more of a list of extent of psychopathology rather than a comment on the mental health of the young person. To that end we wondered about including an explanatory paragraph at the very beginning of the document emphasising the importance of mental health (in Victoria this would be consistent with the recovery model?) I'm sure that you've looked at the 0 to 3 DC 0 to 3 forms and measures of infant mental health and relationship adjustment: http://www.zerotothree.org/childdevelopment/earlychildhoodmentalhealth/dc03rforms.html There will be an announcement about revisions of the 0 to 3R at the forthcoming waimh Congress in Prague. You may know that the new diagnostic classification will be for 0 to 5 years of age roughly and called "DC 0 to 5". The classification will remain a multiaxial one which I think is very helpful for conceptualising the infant in the context of the family and social network. At this stage I'm not sure whether the other axes will be changed substantially or not.

Scale 1: Problems with disruptive behaviour /under controlled emotional regulation	
Include	Oppositional & disruptive behaviours are normal in the development of young children as they explore and begin to individuate. This scale addresses problems with the age and developmentally appropriate capacity of the infant to manage strong feelings, without recourse to age inappropriate levels of overt disruptive behaviours. Clinically, the identification of age-inappropriate emotional regulation does not indicate the source of any difficulties. It may be expected, though not invariable, that regulating emotions connected to hunger, tiredness, and separation may be more prominent with younger infants while overt aggression or rage may be prominent with the older children. Include behaviour associated with any disorder (such as hyperkinetic disorder, depression, autism). Include the capacity to manage intense feelings of hunger, tiredness or separation from the primary caregiver. Include difficulty settling, demanding, whining, undue irritability, excessive crying, frequently arching back and stiffening coupled with turning away from all eye contact, inhibition of expected interactive response, physiological indicators of stress (hiccups, yawns, non-injurious scratching). Include physical or verbal aggression (e.g. pushing, hitting, biting, kicking, teasing), to others (e.g. children, parents or other caregivers, siblings, familiar adults or strangers), animals or objects (e.g. toys). Include oppositional behaviour (e.g. defiance, opposition to authority or tantrums).
Exclude	Problems associated with feeding and sleeping scored at scale 4 (feeding) and scale 8 (sleeping). Problems directly associated with physical health illnesses or disability scored at scale 6. Problems associated with self-injury rated at scale 3.

Rating	Description
0	No problems of this kind during the period rated.
1	Minor problems that are related to a transitional developmental stage, illness or circumscribed change in family circumstances.
2	Mild but definite disruptive behaviour, likely to be limited to one context.
3	Moderately severe disruptive or aggressive behaviour, likely to be in more than one context (e.g. home, childcare, other social situations).
4	Disruptive or aggressive in almost all activities At least one serious physical attack on others or animals, or serious destruction of property.

Additional Responses: Scale 1

- Straight-forward
- Rated with relative ease
- Similar to HoNOSCA
- Very general - feeding/sleeping – emotional dysregulation
    - Not specific enough
    - She is very experienced (gets 3's and 4's a lot)
    - Would not use the scale to see if she would need to do something about
    - Use as assessment marker
      – would not pick up enough individual interactions - causes difficultly
    - Would use for med students
- Title: Problems with Emotional Regulation – do not combine with disruptive behavior/under controlled OR make this scale *disruptive behavior* and put *emotional regulation* into Scale 9.
- Above ratings: have a lead question: *Can you tell me about how your child is managing when they get tired?*
- *turning away from all eye contact* – why is this in this scale? Turning away – over-controlled/excessive control
- *inhibition of expected interactive response* – not usually part of disruptive behavior categories (?elsewhere)
- Rating 1: Minor problems that ~~are~~ could be related to a transitional – to not infer causality
- Rating 4: *serious destruction of property* – formal way of putting it
- Ratings – examples always good [needed]
- Highlighted parental report bias
- How overwhelmed the parent is
- Scale 1 - ticked

Scale 2: Problems with activity levels, joint and/or sustained attention	
Include	Include problems with overactivity/underactivity, joint and sustained attention associated with any cause, including related to aspects of the caregiving environment (e.g. lack of appropriate stimulation, opportunities for motor development). Include problems with restlessness, fidgeting, distractibility, jerkiness or concentration due to any cause, including depression. Include issues of sustained as well as joint attention. Activity and attention difficulties may manifest in altered levels of vigilance, impaired turn taking in behavioural interactions, pronounced startle reflexes and rigidity. Where two factors appear to negate each other (e.g. joint attention problematic but sustained attention highly powerful), rate the most severe occurrence.

Rating	Description
0	No concerns during reporting period.
1	Minor periods of overactivity/underactivity.
2	Some vulnerability in joint and/or sustained attention however the infant's development appears unaffected. Mild overactivity/underactivity or restlessness but with age-appropriate support/structure, the infant can modify their activity levels.
3	Regular significant issues, at times, with overactivity/underactivity that is difficult to control even with appropriate supports. Moderate problems in joint or sustained attention that is typically unable to be modified.
4	Severe levels of overactivity/underactivity impacts negatively on the infant's capacity to engage and achieve developmental milestones across multiple contexts. Consistent and severe limitations in sustained and/or joint attention.

Additional Responses: Scale 2

- Overactivity/underactivity – clarify box statement [rating box] which one is being scored – HoNOSCA fairly unreliable from clinician to clinician
- Joins Scale 11 – Sensitivity links in
- Generally Happy with Scale 2
- Fine
- Related – why ? Specifying a cause. - including ~~related~~ "consequences of"…or…"secondary to" aspects of the caregiving environment.

- Include problems with restlessness, fidgeting, distractibility, jerkiness or concentration *due to any cause* – causality
- *including depression.* Infants – not relevant
- *Where two factors appear to negate each other (e.g. joint attention problematic but sustained attention highly powerful),* - not negating, more contract or other – negate is wrong word
- Rating 2: *vulnerability* – what to make of that "inability to sustain"
- Rating 3: *Regular* means specific intervals. I guess they mean "frequent"
- Rating 4: *Consistent and severe limitations* – good; Change rating 2 to "some limitations"
- Overactivity/Underactivity – too polarized in the Scale – does not account for fluctuations
- All right
- Title: *joint and/or sustained attention* - ?whether should be two separate scales (need to capture people most impaired
- *Jerkiness* - circled
- *Activity and attention difficulties may manifest in altered levels of vigilance, impaired turn taking in behavioural interactions, pronounced startle reflexes and rigidity.* - ?Arousal psychological

Scale 3 Non accidental self-injury	
Rating	**Description**
Include	With infants and pre-schoolers, the question of intentionality is less clear than with older children. While intention should be considered, it will not always be apparent and the clinician may draw on clinical experience to infer intentionality. Behaviours included here are essentially those that result in self-harm that are not the consequence of an accident. However, self-injurious behaviours and actions are rated here irrespective of any indication of intent. Self-soothing behaviour that results in injury or harm e.g. hitting, biting, hair pulling, head banging, rocking, cutting, scratching, excessive sucking leaving marks. May include lack of self-protective reflexes, inhibition of pain and reassurance responses e.g. when an infant is clearly hurt yet inhibits a response where other infants of the same age would be expected to cry, flinch and look to parent(s) for reassurance. Include pretend play involving self-injury such as running into hard objects deliberately in the context of play, making dolls hit their heads during play. Include attempts to stab self with a pen or other non-lethal object, cutting self with knives or scissors, deliberately jumping from a height with injurious intent, frequently discussing intent to self-injure.
Exclude	Self-injurious behaviour secondary to a medical condition. Accidental self-injury unless clearly from a lack of self-protective reflexes.

Rating	Description
0	No problems in the rating period.
1	Occasional occurrence of minor lack of self-protective reflexes, thoughts about self-injury, pretend play involving self-injury, deliberate physical recklessness.
2	Occasional actual self-injury including actions such as rubbing, scratching or rocking which lead to mild levels of physical injury. Pretend play that regularly involves self-injury or self-soothing behaviours that result in harm. Occasional episodes where self-protective reflex is inhibited.
3	Moderately severe occasions of potentially or actually physically significant self-injury. Preoccupation with self-injury. Repeated episodes of self injurious behaviour. In younger infants – inhibition of response to pain/discomfort and lack of self-protection and self-soothing.
4	Severe or regular self-injury occurs. Episodes of physically significant self-injury. Overt inhibition of response to pain/discomfort and lack of self-protection and self-soothing.

Additional Responses: Scale 3

- *Making dolls hit their heads during play* - query of whether it's the child hitting the doll to their head OR hitting the dolls' heads together – self-injury to the play object. Also banging cars could be considered.
- Would not see this group often – only 3's and 4's . Usually a whole lot of other people involved in that
- Referring to Title: "Funny term"
- *Running into hard objects* – is more self-regulation/arousal levels
- *Include attempts to stab self with a pen or other non-lethal object* – not 0-4 years olds. For older children
- *Accidental self-injury unless clearly from a lack of self-protective reflexes.*2 groups: (1) self-soothing, (2) risk-taking
- Core things are OK
- *Self-soothing behaviour …scratching* - Scale 1 mentions non-injurious scratching – if defined previously as non-injury, can't include here
- Title: *accidental* – what is defined as accidental?
- *Accidental self-injury unless clearly from a lack of self-protective reflexes.* – vague
- *What is normal level of injury during age group of learning to walk, exploring? Include pretend play involving self-injury such as running into hard objects deliberately in the context of play, making dolls hit their heads during play.* ?play +destructive play + injury as being different
- Rating 2: "…harm" to who and how is harm defined?

Scale 4: Problems with feeding behaviour	
Include	Feeding behaviours progress with development. The acknowledgement of problems in this area will be influenced by the duration, distress and incongruence of the concerning behaviours with the infant's age and age appropriate development. All glossary examples should be developmentally referenced. Include problems related to difficulties with breast feeding, bottle feeding and solids. Include all feeding difficulties irrespective of potential cause or solution. Include behaviours such as reluctance, resistance or refusing to feed; tiring or sleeping readily during feeding; feeding related distress (e.g. fussiness or crying); maintaining adequate nutrition which may result in nasogastric / gastrostomy tube feedings; sensory adversity; vomiting and difficulty in achieving developmentally appropriate food or feeding skills e.g. limited diet, consistent refusal of certain foods, groups, or types (e.g. solids), or modes of eating; little recognition of the relationship between hunger, feeding and satiety. Include feeding problems related to prematurity, physiological problems and gastrointestinal symptoms.

Rating	Description
0	No problems of this kind during the period rated.
1	Slight occasional problems only. Problems may be transient though typical for the infant's developmental stage.
2	Mild but definite problem with non-organic somatic symptoms. May require additional support to feed. Nutritional intake and growth are likely to be within expected parameters.
3	Moderately severe feeding problems. Some risk of nutritional problems.
4	Severe problems with feeding. May include a high probability of severe nutritional problems.

Additional Responses: Scale 4

- Self-explanatory
- Useful in frame of thinking from birth around feeding
- Sees 3's and 4's. Feeding difficulties well established

- *Mild but definite problem with non-organic somatic symptoms.* I don't know…what does that mean? You don't have to know what the cause is – categorise organic/non-organic. "Organic" versus "non-organic" – not used anymore. Add: "only mild but definite problem"
- Hard to say child has nutritional problems eg OT vs nutritionist
- Rating 1 *typical* implies it happens routinely – change to "uncommonly"
- *Problems may be transient though typical for the infant's developmental stage*. - confusing
- Rating 2 first line "non-organic symptoms"
    - Don't know what these are
    - Need paediatrician's report
    - Do not like "non-organic"
- *Include behaviours such as reluctance, resistance or refusing to feed; tiring or sleeping readily during feeding; feeding related distress (e.g. fussiness or crying); maintaining adequate nutrition which may result in nasogastric / gastrostomy tube feedings; sensory adversity; vomiting and difficulty in achieving developmentally appropriate food or feeding skills e.g. limited diet, consistent refusal of certain foods, groups, or types (e.g. solids), or modes of eating; little recognition of the relationship between hunger, feeding and satiety.*
    - – weight should be accounted for
- *Include feeding problems related to prematurity, physiological problems and gastrointestinal symptoms.* – failure to thrive; percentile chart
- Contextual :
    - How much parents cater for fussiness?
    - Vegemite sandwiches or eat everything?
    - Cultural expectation – finish everything in bowl – by 3 [years], feed themselves
- Include box: independent eating for the older child?

Scale 5: Problems with language or communication skills or specific developmental delays.	
Include	Include cognitive, motor, language, and communication delays irrespective of putative cause or solution. Problems in speech or language associated with any disorder or problem, such as a specific developmental learning problem. Concerns should be rated irrespective of whether additional professional assessment or intervention has occurred (e.g. paediatrics, speech pathology). Language or communication difficulties may include articulation, pragmatics, gestures, vocal quality or range, echoing, interference with vocalisation (e.g. dummy, fingers). Difficulties in this area may impact on ability to communicate and form relationships with similar aged peers and adults, emotional dysregulation, and behavioural issues in the context of difficulty in communication.
Exclude	Physical illness or disability problems such as vision and hearing problems (rated at scale 6). Do not include infants with global learning disability (Intellectual Disability) unless their functioning is below the expected level for them. There will be periods in the very young where the presence of a global learning disability is not conclusively established. It may be expected that problems may be rated on this scale that will subsequently be no longer rated.

Rating	**Description**
0	No language, communication or specific developmental delay issues identified during the rating period.
1	Some minor concerns noted by parent or clinician in regards to delay in speech and language development or developmental delays that may be monitored but likely to be within the normal range of development.
2	Mildly severe concerns noted across more than one setting and in comparison to peers of a similar age.
3	Moderately severe and/ or multiple concerns noted across settings compared with peers of a similar age.
4	Severe problems of language, communication or development exist. Likely to cause significant distress for the infant and/or family. Severe delays compared to peers of a similar age.

Additional Responses: Scale 5

- *Global learning disability* - same dilemma re HoNOSCA – difficult to assess unless you see a lot of children with intellectual delay, hard to say comparison – child with hearing problems – to exclude that? – would have speech delay due to hearing loss . Tendency would be to include even though it says exclude

- Language versus sensory as an assessment tool [need to distinguish]
- Over- or under- mix of those is difficult – speech pathology
- Language or communication skills – under 3 years – attachment related in this setting (see's 3's and 4's)
- *Language or communication skills* combined with *specific developmental delays* but what does *Specific developmental delays* mean? Could be with Scale 6. Motor = cognitive
- Ratings: mildly severe, moderately severe and severe Vs nothing, slight, etc re categories for severity. Why? Speech therapist – developmental milestones for speech and language not well defined. Do they really know? Need notes down the bottom assuming, MDT have been made.
- Rating 1 – change "likely" to "may be" – to say its likely in the rapid tick-box method would be tricky
- *Within the normal range of development* – multiple backgrounds
- Rating 2 – mildly severe - not a clue what that means. If I've got to rate fast, don't know what I'd do.
- Paediatrician, OT, Speech Pathologist; hard to ascertain validity
- Title: *or specific developmental delays* – why add this, then exclude global, when it's a developmental delay?
- Do not include infants with global learning disability (Intellectual Disability) – Why?, when none of the others exclude (specifically) intellectual disability
- Ratings 0,1,2 – can be vague in terms of judging. 3,4 = pretty clear
- Really difficult one because of range:
    - Birth order
    - First-born talks much more than subsequent kids
    - Boys/girls differ
- Even in the same pre-school class, have range of development
- Multilingual - proven to delay language development
- Include box: doesn't include comprehension , sentence structure, vocabulary as a language or communication difficulty
- Exclude box: functioning – unclear of what that means

Scale 6: Problems with physical illness or disability	
Include	Physical health problem or disability which limits or prevents movement, impairs sight or hearing or otherwise interferes with functioning. Problems in this area may be observed or based on reports from others. Movement disorder, side effects from medication. Congenital health problems. Physical health problems that are the result of psychological issues, deprivation, maltreatment or self-injurious behaviour. Include: side effects from medication, physical effects from drug/alcohol use, or physical complications of psychological disorders such as severe weight loss. Include physical complications or disability as consequence of self injury. Ratings will be influenced by consideration of impact of illness on everyday functioning.

Rating	Description
0	No incapacity as a result of physical health problems during the period rated.
1	Slight incapacity as the result of a health problem or disability during the period (e.g. cold, non-serious fall, teething). Parent voices concern about transient physical illness or physical symptoms but these are not considered serious by the parent or clinician.
2	The infant has some mild symptoms of physical illness or disability, that occasionally prevent engagement in usual activities, however the overall structure of their day is preserved and the ability to play is only mildly affected.
3	The infant is suffering moderate symptoms of physical illness or disability, resulting in some ongoing distress and loss of function. There is still some time each day, in which the infant/child is able to engage in usual activities, including play.
4	The infant is suffering severe symptoms of physical illness or disability that result in serious distress and/or loss of function. Normal everyday routines and activities, including play, are generally not possible because of the physical problem. Considerable input of effort and resources may be required to care for the infant, and support the parent.

Additional Responses: Scale 6

- No issue with this one
- Sees severe cases – around eating or relationship with mother – usually 4
- Look up ARACY "The Common Approach" re representing child domains – child, family, community – to guide the different scales. Examples in all domains (there are 6). Examples in each of those issues eg -

family - goes to preschool. Support questions. Eg are there struggles with each domain. Prompt questions to explore sections.

- *Physical health problems that are the result of psychological issues,* - reminds me of non-organic symptoms – requires expertise in that area
- *side effects from medication, physical effects from drug/alcohol use,* have not met a 4 year old who is using drugs or alcohol. ?foetal alcohol spectrum disorder – it's a big deal now – they haven't mentioned it
- Parent voices concern about transient physical illness or physical symptoms but these are not considered serious by the parent or clinician. – seems contradictory
- Easier to describe severe rather than mild – all scales. Hard to define borderline.
- *Include: side effects from medication, physical effects from drug/alcohol use, or physical complications of psychological disorders such as severe weight loss.*- transient illness
- Eg gastro in the last 2 weeks will impact the rating of these
- Sounds like they are trying to infer long-standing organic illness – oh I see cold etc., so you are allowed to rate it
- Include: transient illness in *Includes*

Scale 7: Problems associated with regulation and integration of sensory processing	
Include	Problems associated with processing, regulating and integrating information from sensory stimuli which interfere with the sensory regulation required for adaptive interaction with and exploration of the world. While problems with sensory organs are rated at scale 6, this scale is more concerned with the processing of otherwise apparently intact sensory organs. For example, the question is not whether their hearing is OK, but rather, is their processing of auditory information peculiar or problematic? Problems associated with sensory processing can reflect hypersensitivity (over-reactive therefore avoidant or fearful/cautious) and / or hyposensitivity (under reactive therefore seeking or impulsive) to one or more normal sensory stimuli. Sensory stimuli include vision, touch, hearing, taste, smell, and spatial awareness including the sensation of movement and awareness of body position in space. Problems associated with the regulation and integration of sensory processing are predictable and usually occur across multiple settings and within multiple relationships. Intensity, frequency, duration and location of problematic sensory stimuli may impact on the infant's presentation. Examples of the manifestation of sensory regulation difficulties include responsiveness to fabrics, movement, travel, focus on apparently irrelevant objects, avoidance of play, abnormal motor tone, floppiness in interactions with parents, frozen watchfulness. They may appear to have a preference for swaddling, or to seeking or avoiding certain fabrics. At more severe levels, the infant may be underweight and presents with failure to thrive due to discomfort and distress of feeding. Problems with sensory regulation may manifest as emotional dysregulation or emergent developmental delays.
Exclude	Problems with physical illness or disability rated at scale 6. Problems associated with specific developmental disabilities rated at scale 5. Problems with disruptive behaviour /under controlled emotional regulation rated at scale 1, while problems with anxiety and depression rated at scale 9. Activity levels, joint and/or sustained attention not appropriate to developmental age rated in scale 2.

Rating	**Description**
0	No sensory processing problems identified during the rating period.
1	Some concerns about sensory processing problems (over or under responding to normal sensory stimuli) however the impact on adaptive daily functioning and exploration of the world is minor.

2	One or more sensory processing problems identified and are impacting the infant. The infant/ and /or family may be showing signs of distress but maintaining appropriate developmental milestones. May become agitated, distressed, or disengaged when exposed to specific sensory
	stimuli. Definite and minor impact on functioning in daily tasks or in on maintaining interactions in primary care-giving relationships.
3	Sensory processing problems are impacting on the infant capacity to engage with the environment. May manifest as diminished exploration and play. Definite and moderate impact on daily functioning.
4	Severe and/or ongoing difficulties related to sensory processing problems directly impacting the infant social, emotional and physical wellbeing. Definite and severe impact.

Additional Responses: Scale 7

- Make clear this is NOT emotional [refers to sensory processing in title]
- Suggested title change to: *Problems associated with regulation of sensory processing and integration and regulation*
- Glad to see that Scale is in – important area that is not always attended to sufficiently.
- But there is overlap. Tactile – difficult to say whether its sensory or behavioural.
- OK. Fine.
- How do you tease out *Sensory Processing* [in title] from *emotional regulation* [in exclusions]
- *specific developmental delays*- does it mean without autism? for egs, no problem, rating = 0
- Not necessarily related to sensory issues:
  Examples of the manifestation of sensory regulation difficulties include responsiveness to fabrics, movement, travel, focus on apparently irrelevant objects, avoidance of play, abnormal motor tone, floppiness in interactions with parents, frozen watchfulness. They may appear to have a preference for swaddling, or to seeking or avoiding certain fabrics. At more severe levels, the infant may be underweight

  Feeding [section above]

  and presents with failure to thrive due to discomfort and distress of feeding. Problems with sensory regulation may manifest as emotional dysregulation or emergent developmental delays.

  [this section]: emotional dysregulation – probably OK

- Exclude: Problems associated with specific developmental disabilities rated at scale 5.

  Problems with disruptive behaviour /under controlled emotional regulation rated at scale 1, while problems with anxiety and depression rated at scale 9.

Scale 8: Problems associated with sleep	
Include	Sleep disturbance is common for infants. Include difficulties in both settling and maintaining sleep irrespective of where the locus of the difficulty is thought to be (infant, parent, living arrangements). Include excessive sleep (e.g. which interferes with opportunities for skills or social development), insufficient sleep (e.g. periods of awakenings or reduced sleep time) and disturbed sleeping (e.g. sleep talking, sleep walking, night terrors, or any other disturbance during sleep when the infant does not seem to respond to the parents). Include snoring or loud mouth breathing with breath holding or gasping.

Rating	Description
0	No problems of this kind during the period rated.
1	Minor problems, typically within expected developmental norms, infrequent and where the family appear to have some approaches that successfully address the problem.
2	Mild problems which are intermittent and the family appear to have some success in addressing the problem.
3	Moderately severe sleep problems. The infant's sleeping pattern is a cause for distress in the parents and family. The sleep disturbance is present most of the time, and may be significantly out of keeping with age expectations.
4	Severe sleep problems. The sleeping pattern is a cause for great distress in the parents and family and may be significantly out of keeping with age norms. The sleep disturbance is present nearly all the time and significantly interferes with development.

Additional Responses: Scale 8

- Self-explanatory
- Mild/minor – she is not involved in these cases
- Should be Scale 1
- So much first voiced by parents
- Biased toward younger infants, not older infants, but is first and foremost on the minds of new parents
- *Snoring* – is a physical problem – why is it categorised as a sleep problem? Significant physical problem – obstructive breathing during sleep is a significant physical issue
- All right

Scale 9: Problems with emotional and related symptoms	
Include	Symptoms of depression, anxiety and phobias. Problems with negative or inhibited affect in the infant suggestive of low mood, anxiety, fear or emotional withdrawal. May include fears, anger, anxiety or emotional withdrawal from parents and others. Include incongruent lack of emotional expression. May be expressed with changes in curiosity, clinging, masking face, incongruent emotional expression, startle responses, crying, anger, hypnotic gaze, withdrawal, and blank expression, exaggerated positive or negative emotional responses. May include excessive stillness, frozen watchfulness, quiet rage and restrictions in affect range. Apparent increased tolerance for aversive adult behaviour should be considered. Include emotional responsiveness to challenges in the environment and requirement for comfort. Include age or developmentally inappropriate lack of wariness, or avoidance of parents.
Exclude	Physical sequelae of psychological disorders or medication – rated at scale 6. Disruptive behaviours resulting from emotional distress – rated at scale 1. The emotion associated with the disruptive behaviour is rated here at scale 9.

Rating	Description
0	No problems in the rated period.
1	Transient and/or minor mood, anxiety and emotional symptoms or changes.
2	Mild but definite emotional symptoms are present but not preoccupying.
3	Moderately severe emotional symptoms are present.
4	Severe emotional symptoms which intrude into all activities and are nearly always uncontrollable.

Additional Responses: Scale 9

- Add at the end of title in brackets "(anxiety exhibitor)" to denote that its looking at the aspect, not the spectrum
- Very focused on inhibition – trying to remember where is there a disinhibition place
    - Emotional dysregulation -a cliché in our line of work
- *Problems with negative or inhibited affect* – turning away Scale 1 – belongs here not Scale 1
- withdrawal from parents and *others* – others too general – change to familiar/significant others – too general to say others
- *exaggerated positive or negative emotional responses.* Exaggerated positive – delineate from negative. Trying to say exaggerating is a problem

- *Physical sequelae of psychological disorders* – how to identify whether in Scale 1 or not. *Disruptive behaviours resulting from emotional distress – rated at scale 1.* Should be here [Scale 9] – turning away – therefore not disruptive
- Scale 9 description – overwhelming to read – a mess of unrelated emotional expressions just listed without any way of structuring them.
- Rating 2: *preoccupying* – what is this trying to say in rating a 1 year old?
- Does not talk about different settings, whereas other descriptions often driven by circumstances/settings – need consistency
- *Symptoms of depression, anxiety and phobias.* – delete
- *Problems with negative or inhibited affect in the infant suggestive of low mood* (depression)*, anxiety, fear/*phobias *or emotional withdrawal.*
- *Include incongruent lack of emotional expression.* – delete
- This *may be expressed*
- Start new paragraph before the word *Apparent*
- Inclusion suggestions:
    - This may include
        - Changes in curiosity…emotional responses
        - Excessive stillness…restrictions in affect range
        - Increase tolerance for aversive adult behavior
        - Emotional responsiveness…comfort
        - Age or developmentally…parents

Scale 10: Problems with social reciprocity	
Include	This scale addresses the capacity of the infant to engage in, and engage others in, age and developmentally appropriate interactions. There may be problems with seeking, engaging and enjoying interactions with familiar adults and children, including development of the social smile at 6 weeks. Responses to social engagement or social intrusion from others are not responded to appropriately e.g. ambiguous half smiles. Problems may manifest in reciprocity in communication, play, and games. Reciprocity may be expressed both pre verbally and verbally, as well as behaviourally. Problems may manifest as indiscriminate and overfamiliar social interactions as well as withdrawn and disengaged social interactions. The scale includes problems with emotional attunement and misalignment between the infant's needs and the parents' or carers' responses. Problems rated in this scale may include the infant's capacity to manage appropriate eye contact e.g. the infant may not gaze at the parent's face or at interesting object when shown. Problems may include avoidant gaze; no eye contact (but no active avoidance either); brief glances without sustained looking (difficulty gaining and sustaining eye contact); and unfocused eyes. Problems with vocalisations relating to reciprocity of interactions, such as turn taking, engagement attempts, and vocal mirroring may also be relevant indicators of social reciprocity issues.
Exclude	Difficulties with vocalisation separate to the social reciprocity function are rated at Scale 5.

Rating	Description
0	No problems during reporting period.
1	Transient or mild problems in the infant's developing capacity to engage in social relationships.
2	Mild but definite problems in at least one area.
3	Moderate problems in more than one area.
4	Severe issues with social reciprocity. Problems likely to occur in most areas and intrude across most interactions.

Additional Responses: Scale 10

- Emphases are in looking at the relational aspect or not. Eg "absence of referencing to parent, but kid who engages with us"
- Social reciprocity tag - linked to Autism Spectrum Disorder (ASD) but theirs' in attachment - band or say: can't say with 2 or 3 year old which it is – make global statement

- how bad is the eye contact – not whether its there or not
- *Problems may manifest ~~in~~ with reciprocity* – edit to make consistent with title of the scale
- Rating 1 – delete *developing*. Rapid, tick-box approach – just thinking of current capacity is difficult
- Rating 4 – "likely" – why *likely*? Most severe category – supposed to talk of an impression.
- *This scale addresses the capacity of the infant to engage in, and engage others in, age and developmentally appropriate interactions.*

  There may be problems with seeking, engaging and enjoying interactions with familiar adults and children, including development of the social smile at 6 weeks. Responses to social engagement or social intrusion from others are not responded to appropriately e.g. ambiguous half smiles. Problems may manifest in reciprocity in communication, play, and games. Reciprocity may be expressed both pre verbally and verbally, as well as behaviourally. Problems may manifest as indiscriminate and overfamiliar social interactions as well as withdrawn and disengaged social interactions.
  Child focus to here

  The scale includes problems with emotional attunement and misalignment between the infant's needs and the parents' or carers' responses.
  Parent response focus. Does not equal child difficulty in Social Reciprocity (parents may not be attuned – does not mean child has deficit.) All lumped together.

  - Whole description is very broad.
  - Rating 2 - vague

Scale 11: Problems with age appropriate self-care and environmental exploration	
Include	This scale addresses age-appropriate self-care and exploration of the environment. Self-care is likely to be a more prominent consideration with older children. Self-care is likely to include age appropriate levels of assistance with bathing, feeding, dressing, playing etc. Problems with contributing to self-care and environmental exploration often exist due to restrictions in the infants' environment. It is critical to rate the problem and good practice to not assume that the rating indicates that the infant is the sole source of that issue. Include problems with basic activities of self-care such as washing, dressing, toileting. Exploration may include visual, tactile, verbal as well as physical exploration (under or over exploration). Include problems with complex skills such as play, autonomous activities or separating from parents, taking into account the norm for the infant's age and developmental stage. Difficulties may be indicated by regression to an earlier stage of development. Problems with separating from parents when the infant is attending structured socialisation settings (e.g. day care, pre-school) may be rated here although the resulting impact on attendance at socialisation settings should be rated at scale 13. Include poor levels of functioning arising from apparent lack of motivation, mood, environmental restriction or any other issue whether it is considered to arise from the infant or the parents or the environment.
Exclude	Do not include feeding problems rated at scale 4. Do not include sleeping problems rated at scale 8. Do not include lack of opportunities for exercising intact abilities and skills, as might occur in an over-restrictive family rated at scale 12. Do not include the outcome of limited environmental exploration on structured socialisation settings rated at scale 13.

Rating	Description
0	No problems during reporting period.
1	Minor problems with self-care or exploration of the environment.
2	Mild problems with self-care or exploration of the environment.
3	Moderate level of problems with self-care or exploration of the environment.
4	Severe level of problems with self-care or exploration of the environment that is likely to be intruding across settings, activities and persons.

Additional Responses: Scale 11

- *Self-care* and *environmental* – what is it that puts this in the same scale?
- If having discussion on attachment theory, they belong together, but as an observation, they do not belong together
- See 0-3, less than 2 years
- *Problems with contributing to self-care and environmental exploration often exist due to restrictions in the infants' environment.* – sounds like an over-restrictive family but in exclusions "Do not include"…restrictive – excluded – confusing to read
- Rating 4 – *likely* – why is this here?
- Contradictory: Include poor levels of functioning arising from apparent lack of motivation, mood, environmental restriction or any other issue whether it is considered to arise from the infant or the parents or the environment. [in exclusions:] Do not include lack of opportunities
- Environmental Restrictions – include or not?

Scale 12: Problems with family life and relationships	
Include	This scale addresses problems in family life that are thought to impact on the infant. If the parents are separated, consider the relationship with each parent and the separated parents' ability to co-parent where appropriate. Include relationships with significant others should be included – grandparents, siblings, extended family members, child care providers. Include instances of neglect including physical (e.g. lack of sufficient access to appropriate food, shelter & clothing) and emotional (e.g. lack of warmth, comfort and age appropriate regulation of the infant's affect). The availability of access to caring attentive and empathic adults, and the ability to keep the infant in mind, should be considered. Include parent or family irritability with the infant. Difficulties in managing powerful emotions or any consequent harmful behaviour by those in the infant's immediate environment should be considered. Include instances of physical or verbal hostility or abuse towards the infant, as well as family hostility or conflict which impacts on the infant. Consider capacity for significant others to contain powerful negative emotions towards the infant. Issues such as parental or sibling mental health, substance use and personality problems should be included if they have an effect on the infant.
Exclude	Do not include disruptive behaviour by infant, rated at scale 1. Do not include problems with social reciprocity rated at scale 10.

Rating	Description
0	No problems during the period rated.
1	Slight or transient minor problems. Some concerns about family relationships are evident but these are part of a history where reasonable parental reflective capacity is typically apparent.
2	Mild but definite problem that have some impact on the infant's development.
3	Moderate problems in the relationship between the infant and primary caregivers or family that are affecting the infant's regulatory and exploratory capacity.
4	Serious problems in family relationships with severe impact on the infant.

Additional Responses: Scale 12

- Referring to title: 4-year-old = safety concerns, often wording is too safe, subtle, soft - versus level of concern you can have for infants at this age. Draw more attention to it.

- *Do not include disruptive behaviour by infant, rated at scale 1.* - Not sure how separate that can be to parents' response to the infant, or put in description
- Referring to Rating Description: If kid is difficult, but family is managing, do not score. Once the family is not managing, family response is problematic – becomes [Scale] 12. Distinction needs thrashing out.
- Description should be "regardless of cause"
- Suggested change to rating 4 to add at the end "/dangerous/unsafe parenting" – with risk to the infants' safety – or Scale 12.
- Struggle when use rating scales for different purposes eg for safely, is inadequate
- If for family – 4 – does not seem to adequately capture the mood in the middle of the interview.
- Warped by the time they get to her [she sees severe cases]
- *availability of access to caring, attentive and empathic* - add comma after *caring*
- This is a good one. [Scale]
- Ratings :
    - Rating 2: broad – some impact - development
    - Rating 3 – narrow - -*affecting the infant's regulatory and exploratory capacity.* Why restricted to this when Ratings 2 and 4 are broad and general?
    - Rating 4 – broad – severe impact – no development
- Delete words as follows: *Include relationships with significant others* ~~should be included~~

Scale 13: Problems with attending care, education and socialisation settings	
Include	This scale addresses attendance at the prime structured socialisation setting outside of the family. Include attendance at any type of regular socialisation and care activity at the time of rating e.g. regular care with extended family or formal early childhood education (sometimes called kindergarten or pre-school). Include activities irrespective of location e.g. regular play group sessions at infant's home. Include refusal of, or withdrawal from early childhood education, childcare, play group or similar regular socialisation activity, irrespective of cause. Include limited or minimal opportunities to attend socialisation activities appropriate to the infant's age. Include consideration of additional supports such as reassurance, transitional objects, required to settle the infant in the setting. If early childhood education, childcare etc. is in holiday break, rate the last two weeks of the previous open period. Note: Infants and young children will communicate their reluctance and distress at attending these settings through a range of symptoms. These may include problems in feeding, toileting, eating, playing, communicating and sleeping both at the settings and around the transition time. These symptoms in themselves are likely to be rated at different HoNOSI scales and are not the sole source of rating at this scale. However, it is acknowledged that the reluctance to attend may be conveyed to the clinician through these symptoms. The actual attendance problems are rated at this scale.
Exclude	All behaviours and emotional expressions of problems associated with attendance or separation are rated at their respective scales (e.g. Disruptive at scale 1, Feeding at scale 4, Emotional at scale 9). Absences due to illness of infant or parents requiring them to be absent from the setting. This typically includes medical conditions, such as fevers, contagious illnesses or infections which would be rated at scale 6.

Rating	Description
0	No problems during the period rated. Infant displays age appropriate behaviour on separation from their parents and settles readily when comforted in the environment.
1	Minor problem with attending and may display reluctance for brief periods. Responds with small amount of support additional to that typically required at this age.
2	Mild but definite problem with some sessions missed or refusal to participate in activities when attending.

3	Moderate problem with several days missed during rating period due to infant's reluctance to attend.
4	Severe problem with infant absent for most of the days or sessions during rating period.

Additional Responses: Scale 13

- Happy with [Scale] 13
- Does not arise much in her role
- Fine.
- Absences due to illness of infant or parents requiring them to be absent from the setting. This typically includes medical conditions, such as fevers, contagious illnesses or infections which would be rated at scale 6. – why rating is child is sick and not attending = (0)? Attendance is black and white – probably for the beginning/whole scale. What is the point of the exclusion and what does it mean for that rating?
- Include consideration of *Take into account* additional supports *that are required to settle the infant in the setting* such as reassurance, transitional objects, required to settle the infant in the setting.
- If early childhood education, childcare etc. is in holiday break, rate the last two weeks of the previous open period. Rate last two weeks when child was expected to attend a structured socialization setting.
- Include box: recommendation: too long: Include; attendance problems are rated on the scale. Other symptoms which are associated with problems with attending…are rated in other scales.

Scale 14: Problems with knowledge or understanding about the nature of the infant's difficulties	
Include	Include lack of useful information or understanding available to the parents, caregivers or referrers or support system about the nature of the difficulties. Include problems with capacity or knowledge to understand the infant's difficulties. Include limited or incorrect understanding of the infant's developmental stage and needs. Include misunderstanding, minimising, elaborating or exaggerating the difficulties, impact or distress as well as inaccurate attribution of the infant's difficulties. Include lack of explanation about the difficulty/diagnosis, the cause of the problem or understanding of the prognosis or the impact on the infant. Rating a problem here does not preclude that the service system may revise their understanding of the infant's difficulties. In many ways, problems rated here may indicate a lack of congruence between the parent's and other key figure's views about the nature of the difficulties and the assessing or treating system's views.

Rating	Description
0	No problems during the period rated. Parents demonstrate a good level of understanding about the difficulties.
1	Minor problems only. For example, parents/ essentially understand infant's difficulties but with occasional misunderstandings such as sometimes downplaying, or exaggerating the infant's difficulty or distress.
2	Mild but definite problem in understanding infant's difficulties.
3	Moderately severe problems. Parents have very little or very poor knowledge about the nature of their infants' problems.
4	Very severe problem. For example, parents have no understanding about the nature of their infant's problems.

Additional Responses: Scale 14

- Very similar to HoNOSCA. Fine.
- *Include misunderstanding, minimising, elaborating* – "elaborating" means explaining in detail, a positive thing. I would remove the word "elaborating"
- Rating 1 – delete the space after "*parents/* " – an edit
- Rating 1: Should *Reflective capacity* be here (is in Scale 12) – does not really matter or change anything
- *Include problems with capacity or knowledge to understand the infant's difficulties.* Change to: Include problems with capacity to understand or knowledge to appreciate the infants' difficulties.

Scale 15: Problems with lack of information, understanding about services, or managing the infant's difficulties	
Include	Include lack of useful information available to the parents, caregivers, or referrers, or a lack of understanding regarding services or management of the difficulties. Include lack of useful information or lack of understanding about the most appropriate social, therapeutic, developmental, care, and educational supports and interventions for the infant's presentation. Include here supports and interventions that the parents could be providing directly to their infant. Include parental willingness to access/utilise services and interventions to support the infant. The consistency with which parent's understand or use appropriate management strategies, and the extent to which supports are required to help the parent's use optimal approaches may be considered here. Include information not acted upon such that appropriate and feasible interventions are not implemented adequately. Include family's unwillingness or social isolation if it restricts engagement with appropriate services. Child protection authorities may be involved if the match between lack of engagement with the provided information and the infant's difficulties are severe enough. Note: The lack of useful information may stem from a failure of the environment to provide such information or from the parent's misunderstanding or unwillingness to act on available information. Include parent's reflective function regarding their role in supporting the infant and their ability to engage with and utilise services available. Rating a problem here does not preclude that the service system may revise their understanding of the optimal approach to managing the infant's difficulties. In many ways, problems rated here may indicate a lack of congruence between the parent's and other key figure's views about the management of the infant's difficulties and the assessing or treating system's views.

Rating	Description
0	No problems during the period rated.
1	Minor problems only. For example, parents have a good enough understanding of how best they and other resources can help their infant, or they are actively seeking appropriate information, support or access to services.
2	Mild but definite problem in understanding the appropriate services, approaches, resources and supports for the infant's difficulties.
3	Moderately severe problems.
4	Very severe problem.

Additional Responses: Scale 15

- Happy
- First line: *Include lack of useful information available to the parents, - service letting parents down but in Ratings, not included* – therefore include an example of a service problem alongside a parent problem.
- *required to help the parent's use optimal* – no apostrophe in parents
- Not clear difference between Scales 14,15

Additional Responses to Overarching Question

- Have intention to use
- Thought is does what it set out to do – be a symptomatic measure rather than a clinical formulation
- Usefullness – HONOS – HoNOSCA more useful – HoNOSI more useful than that but sees no value in any of them, too clumsy, not clinically useful. What clinician does in practice is ask family, make observation, intervening in the run (clinical formulation) – this [measure] could be an add-on rather than a clinical tool. Does not base a case on a score – symptom focused issue. Symptoms versus outcome – tends to be more global. "I fill in forms because my admin requires me to fill in forms."
- [another person, same group] As a clinician, – helps us be aware/notice – brings back to consciousness things we are working with.
- Yes. Effective at Childrens' services
- Wouldn't be reaching for this to help her make decisions
- More useful down one level in Health Care Services (this is tertiary)
- Good its being made
- "Really useful thing". Not necessarily going to assist in tertiary setting unless in MDT. Uses much more specific scales et on feeding
- Down a level – absolutely – Childrens' centred, very useful, at generalistic presentation eg 2 or 3 or 1 [ratings] – Child and Youth Health
- I think it is, yes. I am quite impressed. For me to read it, it was quite a useful trigger to provide a framework. I liked:
    - Framework
    - Flow
    - Most of the categories
- Unusual – lack of information/understanding about services – surprised to see them there, but they make sense.
- Does – examples, think of specific examples, can see how any changes can be monitored over a period of time
- How long is it meant to take?
- 0-3 years – massive developmental change

Annex 2: Review of Infant Outcome Measures Report

Review of Infant Outcome Measures Report

CONTENTS

Australian Mental Health Outcomes and Classification Network
Review of Infant Mental Health Outcome Measures

Background

Routine Outcome Measures embedded in mental health service delivery are used at all levels; from policy-makers and planners to inform decisions regarding system-wide reforms, to service managers monitoring quality and effectiveness, to clinicians consulting with consumers (Trauer, 2010). In late 2003, a consortium contracted under the Australian Government's Department of Health and Aging (currently the Department of Health) was formed to embed routine outcome measurement into service delivery and to manage, analyse and report on these large volumes of outcome data (Burgess, Pirkis and Coombs, 2015).

To advise on routine outcome measurement from the perspective of clinicians, consumers and carers, three expert groups were formed in 2004 (Children/adolescents; Adults; Older Adults) and in close collaboration with the Australian Mental Health Outcomes and Classification Network (AMHOCN)(Australian Mental Health Outcomes and Classification Network, 2016) this developed into the National Mental Health Information Development Expert Advisory Panel, along with a fourth panel (Forensic). These panels have advised the Australian Government at all stages of implementation.

In Australia, consumer outcome measurement is collected under the National Outcomes and Casemix Collection (NOCC). Following literature reviews, consultation with stakeholders and empirical field trials, under the NOCC Protocol, the Health of the Nation Outcomes Scales for Children and Adolescents (HoNOSCA) is used in mental health services for children and adolescents (Burgess, Pirkis and Coombs, 2015). The NOCC Protocol also specifies which measures are to be used at older stages in the lifespan; Adults and Older Adults. These measures are useful for clinical practice, service delivery and workforce development (Burgess, Pirkis and Coombs, 2015). However, the NOCC protocol does not specify an outcome measure for use with infants or pre-schoolers. There was a need to investigate the availability and feasibility of measures for infants to provide outcome measurement for the entire lifespan. To that end, AMHOCN commissioned a review of outcome measures for infants. Unless further specified, 'infants' will be used in this paper to refer to all those aged under 4 years old.

In addition to the specified instruments, the NOCC protocol attends to the source of the information and generally attempts to include clinician based, consumer based and carer based instruments. The instruments are also required to be broad enough to apply to the range of mental health problems presented in the public and private mental health systems (Australian Mental Health Outcomes and Classification Network, 2012).

Aim

The aim of this study was to investigate candidate measures for infants and pre-schoolers within the parameters of the National Mental Health Outcomes and Classification Framework. With specific reference to the target population, infant social and emotional functioning is strongly influenced by the domains of parent-infant relationships, parent-family factors and by infant health and development (ZERO to THREE. Early connections last a lifetime, 2018). This review aimed to update previous work undertaken to identify suitable routine outcome measures for infants; identify new information on the psychometric properties of the previously highlighted instruments (Child and Adolescent Mental Health Outcomes Expert Advisory Panel, 2008) and to identify which psychometric tools can be used under which circumstances. This review was designed to adhere to the criteria for outcome measures used

Methods

A comprehensive literature review was undertaken in order to investigate the suitability of currently available measures. This review was performed in three stages. First, Medline and PsychInfo databases, including exploration of MeSH subject headings, were searched. The Medline search (1946 to March 2012) was performed using a keyword search utilizing Boolean operators:

> (psych* and (status or rating or scale or test or instrument or questionnaire or measure or psychometrics)) and (social and (behaviour or behavior)) and disorder and infant

This yielded 56 results. The same search was then performed on PsychInfo (1806 to March Week 3 2012) and revealed two further relevant references. As the focus of the review was on sensitivity to change and applicability across the broad range of mental health presentations, psychometrics with a primary focus of screening, or papers that focused on a particular problem such as autism, were excluded.

Secondly, the above results were examined for additional references and retrieved. The output of the above steps were cross-validated against the previous review (Child and Adolescent Mental Health Outcomes Expert Advisory Panel, 2008) and background working papers (Kowalenko and Reilly, 2007). Finally, an Internet search using the Google search engine was conducted to investigate any further developments in the field not covered via the methods described above.

Evaluation Criteria Applied to Outcome Measures Identified

A hierarchical criterion-based approach was used to evaluate candidate outcome measures. While the criterion was hierarchical, many instruments partially met a criterion (eg. part of, but not all of the age range). Instruments which met the criteria partially in this way, remained, then were reassessed for inclusion or exclusion at the next step and so forth for each of the assessment criteria. The criteria considered and the order in which they were considered, were as follows:

- **Age** - Suitable for use with infants and young children 0-47 months

- **Completion Rate** - Is brief and easy to use (≤50 items) and able to be completed quickly (under 10 mins) by a parent, carer or clinician
- **Social/Emotional/Behavioural Domains** - Explicitly measures social, emotional and behavioural domains of functioning (and changes there-in) related to infants and young children 0-47 months
- **Available** - Is readily available (consideration given to cost, copyright issues)
- **Sound Psychometric Properties demonstrated** (e.g., of internal consistency, validity, reliability and sensitivity to change)
- **Scientific** - Has undergone appropriate processes for developing and piloting and has been scientifically scrutinised
- **Continuity** - Has some continuity with outcome measures used with children at later life stages (HoNOSCA, SDQ) and may include ratings by clinician, consumer (or proxy) or other
- **Quantitative** - Yields quantitative data
- **Generalizable** - Is applicable in the Australian context

Results

The search strategies revealed 18 instruments. The following psychometric tools were considered in the current review.

- **Ages and Stages Questionnaire** and **Ages and Stages Questionnaire: Social-Emotional** (Brookes Publishing Company, 2012a, 2012c; Limbos and Joyce, 2011; Mackrides and Ryherd, 2011; Marks, 2007)
- **Parents' Evaluation of Developmental Stages (PEDS)** (Limbos and Joyce, 2011)
- **ITSEA** (Margaret J. Briggs-Gowan, Carter, Irwin, et al., 2004; M. J. Briggs-Gowan, Carter, Skuban, et al.; Alice S. Carter, Briggs-Gowan, Jones, et al., 2003; A. S. Carter, Garrity-Rokous, Chazan-Cohen, et al., 2001)
- **BITSEA** (Margaret J. Briggs-Gowan and Carter, 2008; Margaret J. Briggs-Gowan, Carter, Irwin, et al., 2004; Haapsamo, Ebeling, Soini, et al., 2009; Karabekiroglu, Briggs-Gowan, Carter, et al., 2010; Kruizinga, Jansen, Carter, et al., 2011)
- **Ghuman-Folstein Screen for Social Interaction** (Ghuman, Freund, Reiss, et al., 1998)
- **HoNOSCA** (Brann, Coleman and Luk, 2001; Eggleston and Watkins, 2008; Gowers, Bailey-Rogers, Shore, et al., 2000; S.G. Gowers, R.C. Harrington, A. Whitton, et al., 1999; Hanssen-Bauer, Aalen, Ruud, et al., 2007; Hanssen-Bauer, Gowers, Aalen, et al., 2007; Kisely, Campbell, Crossman, et al., 2007; Lesinskiene, Senina and Ranceva, 2007)
- **Neonatal Behavioural Assessment Scale** (Girling, 2006)
- **Denver Developmental Screening Test (DDST)** (Smith and Brun, 2006)
- **Child Behavior Checklist** (Margaret J. Briggs-Gowan and Carter, 2008; M. J. Briggs-Gowan, Carter, Skuban, et al.; A. S. Carter, Garrity-Rokous, Chazan-Cohen, et al., 2001; Skovgaard, Olsen, Christiansen, et al., 2008; Smith and Brun, 2006)
- I**nfant Toddler Symptom Checklist** (Skovgaard, Olsen, Christiansen, et al., 2008)
- **Baylay Scales of Infant Toddler Development (BSID III)** (Skovgaard, Olsen, Christiansen, et al., 2008)

- **Baby Alarm Distress Scale** (Mantymaa, Puura, Luoma, et al., 2008)
- **Interaction Rating Scale** (Anme, Shinohara, Sugisawa, et al., 2010)
- **Caregiver-Teacher Report Form (C-TRF 2-5) - Achenbach System of Empirically Based Assessment (ASEBA)** (T. Achenbach, 2012; T. M. Achenbach, Edelbrock and Howell, 1987).
- **Parent-Infant Relationship Global Assessment Scale (PIR-GAS)** (The Washington Institute and the Washington State Mental Health Division, 2012)
- **Parenting-Stress Index (PSI) – Short Form** (American Psychological Association, 2012)
- **Strengths and Difficulties Questionnaire** (youthinmind, 2012) Goodman, Robert
- **PSC-17** (The California Evidence-Based Clearinghouse for Child Welfare, 2012)

Each of the above listed instruments were assessed against the criteria and the results are summarised in the tables below. Further details and sources can be found in Appendix 1: Summary of Psychometric Tools Assessed against Evaluation Criteria. When formulating recommendations for psychometric tools, the various domains need also be considered (Australian Mental Health Outcomes and Classification Network, 2012). The domains considered are:

- Clinician-based
- Consumer
- General Functioning

Criterion 1: Age - Suitable for use with infants and young children 0-47 months

As can be seen from Table 1 below, 6 of the 18 measures explored fit the criteria of being suitable for 0-47 months in age. Of those, 2 are suitable from 3 months of age (ASQ:SE and PSI/SF). It was unclear whether the Interaction Rating Scale was suitable for <18 month olds. Measures which did not cover the full age range of 0-47 months such as the ITSEA, BITSEA, SSI. HoNOSCA, etc., were considered as possibility useful for part of this age range (eg. 12-35 months with BITSEA) and therefore included for further evaluation against other criteria.

Table1: *Psychometric Tools against Criteria 1 - Age*

Psychometric tool	**Age**	Met Criteria? Yes = ☑ No = ☒ Potential = ?
Ages and Stages Questionnaire: Social-Emotional (ASQ:SE)	3-66 months	☑
Parents' Evaluation of Developmental Stages (PEDS)	0-8 years	☑
Infant-Toddler Social and Emotional Assessment (ITSEA)	12-36 months	?
Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	12-35 months	?
Ghuman-Folstein Screen for Social Interaction (SSI)	2-5 years	?
HoNOSCA	3-18 years	?
Neonatal Behavioral Assessment Scale (NBAS)	0-2 months	☒
Denver Developmental Screening Test (DDST)	2 weeks to 6 years	☑
Child Behavior Checklist (CBC)	1 ½ to 5 years	?
Infant Toddler Symptom Checklist (ITSC)	7-30 months only	☒
Bayley Scales of Infant Development (BSID II)	1-42 months	☑
Baby Alarm Distress Scale (BADS)	2-24 months only	?
Interaction Rating Scale (IRS)	Participants in study were 18, 30, 42 months. Unclear whether can be used I <18 months	?
Caregiver-Teacher Report Form (C-TRF 2-5) - Achenbach System of Empirically Based Assessment (ASEBA)	1 ½ to 5 years	?
Parent-Infant Relationship Global Assessment Scale (PIR-GAS)	0-5 years	☑
Parenting-Stress Index – Short Form (PSI/SF)	3 months to 12 years. Primary population is 0-3 years.	☑
Strengths and Difficulties Questionnaire	3-16 years	?
Pediatric Symptom Checklist (PSC-17)	4-18 years	☒

Criterion 2: Completion Rate – Is brief and easy to use (≤50 items) and able to be completed quickly (under 10 mins) by a parent, carer or clinician

Table 2 shows that the majority of the age-appropriate scales fail on the requirement of being brief and easy to use with 50 items or less, and able to be completed in less than 10 minutes.

Table 2: *Psychometric Tools against Criteria 2 – Completion Rate*

Psychometric tool	**Completion Rate** (≤50 items, under 10 mins)	Met Criteria? Yes=☑No=☒
Ages and Stages Questionnaire: Social-Emotional (ASQ:SE)	10-15 mins to complete for social-emotional subscale.	?
Parents' Evaluation of Developmental Stages (PEDS)	10 items, few mins	☑
Infant-Toddler Social and Emotional Assessment (ITSEA)	25-30 mins	☒
Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	7-10 mins, 42 items	☑
Ghuman-Folstein Screen for Social Interaction (SSI)	54 items	☒
HoNOSCA	8 mins on average to complete	☑
Denver Developmental Screening Test (DDST)	20 minutes required	☒
Child Behavior Checklist (CBC)	>50 questions total	☒
Bayley Scales of Infant Development (BSID II)	30-60 mins	☒
Baby Alarm Distress Scale (BADS)	14 items	☑
Interaction Rating Scale (IRS)	Too lengthy. 70 items for behavioural score, plus 11 items for impression score.	☒
Caregiver-Teacher Report Form (C-TRF 2-5) - Achenbach System of Empirically Based Assessment (ASEBA)	99 items	☒
Parent-Infant Relationship Global Assessment Scale (PIR-GAS)	10-point scale	☑
Parenting-Stress Index – Short Form (PSI/SF)	36 items; 10 mins	☑
Strengths and Difficulties Questionnaire	33 questions	☑

Criterion 3: Domains - Explicitly measures social, emotional and behavioural domains of functioning (and changes there-in)**.** The psychometric tools above also measure the domains of interest (Table 3).

Table 3: *Psychometric Tools against Criteria 3 – Domains*

Psychometric tool	**Domain**	Met Criteria? Yes=☑No=☒
Ages and Stages Questionnaire: Social-Emotional (ASQ:SE)	Social/emotional, developmental delays	☑
Parents' Evaluation of Developmental Stages (PEDS)	For developmental and behavioural problems	?
Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	Externalising, Internalising, Dysregulation, Competence	☑
HoNOSCA	Psycho-social, emotional	☑
Baby Alarm Distress Scale (BADS)	Facial expression, eye contact, general activity, self-stimulating gestures, vocalizations, response to stimulation, relationship to observer, ability to attract attention	☑
Parent-Infant Relationship Global Assessment Scale (PIR-GAS)	Parent-Infant relationship	☑
Parenting-Stress Index – Short Form (PSI/SF)	Child subscales: Adaptability, Acceptability, Distractibility/Hyperactivity, Demandingness, Mood, Reinforces Parent	☑
Strengths and Difficulties Questionnaire	Emotional, conduct, hyperactivity/inattention, peer relationship problems, prosocial behaviour	☑

Criterion 4: Available - Is readily available (consideration given to cost, copyright issues)
Other than the HoNOSCA, PIR-GAS and SDQ, all remaining psychometric tools were found to have cost associated with them. This is an issue for large-scale use, subject to negotiation with the authors.

Table 4: *Psychometric Tools against Criteria 4 – availability*

Psychometric tool	**Available**	Met Criteria? Yes=☑No=☒
Ages and Stages Questionnaire: Social-Emotional (ASQ:SE)	Low-cost	☑
Parents' Evaluation of Developmental Stages (PEDS)	$55 for set of 50	☑
Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	$110.25 for set of 25 forms	☑
HoNOSCA	Freely available	☑
Baby Alarm Distress Scale (BADS)	Free	☑
Parent-Infant Relationship Global Assessment Scale (PIR-GAS)	Free	☑
Parenting-Stress Index – Short Form (PSI/SF)	$126.00 for kit with 25 forms	☑
Strengths and Difficulties Questionnaire	Free of charge for non-commercial purposes	☑

Criterion 5: Sound Psychometric Properties

The remaining psychometric tools were found to have sound psychometric properties (Table 5).

Table 5: *Psychometric Tools against Criteria 5 – Psychometric Properties*

Psychometric tool	**Sound Psychometric Properties** (e.g., of internal consistency, validity, reliability and sensitivity to change)	Met Criteria? Yes=☑No=☒
Ages and Stages Questionnaire: Social-Emotional (ASQ:SE)	Test-retest reliability = 0.92, Inter-rater reliability=0.93. Validity=0.82-0.88, Sensitivity = 0.86, Specificity=0.85	☑
Parents' Evaluation of Developmental Stages (PEDS)	-Sensitivity and specificity 70-80% -False positives are high	☑
Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	Test-retest reliability excellent, high internal consistency. Inter-rater reliability statistically significant. Found to be a valid and reliable measure	☑
HoNOSCA	Good psychometric properties. Interclass correlation co-efficient = 0.84	☑
Baby Alarm Distress Scale (BADS)	Good validity, sensitivity = 0.82 and specificity =0.78. Good construct validity, satisfactory reliability.	☑
Parent-Infant Relationship Global Assessment Scale (PIR-GAS)	Predictive validity demonstrated	☑
Parenting-Stress Index – Short Form (PSI/SF)	Very good reliability parent =0.55-0.80; child 0.62-0.70. Test-retest reliability after 1 year: 0.70 Parent, 0.55 Child; After 3 weeks: 0.71 Parent and 0.82 Child. Validity: 41% on child section accounted for by 6 factors, 44% on parent section accounted for by 7 factors.	☑
Strengths and Difficulties Questionnaire	Psychometric properties robustly tested and validated	☑

Criterion 6: Scientific - Has undergone appropriate processes for developing and piloting and have been scientifically scrutinised

These tools were also found to have undergone scientific scrutiny by evidence of publication in peer reviewed literature, quantitative and are able to be used in the general population (Table 6). Of those, the HoNOSCA, PIR-GAS and SDQ provide continuity.

Table 6: *Psychometric Tools against Criteria 6 – Scientific*

Psychometric tool Met Criteria? Yes=☑No=☒	**Scientific**	**Continuity**	**Quantitative**	**Generalisable**
Ages and Stages Questionnaire: Social-Emotional (ASQ:SE)	☑	☒	☑	☑
Parents' Evaluation of Developmental Stages (PEDS)	☑	☒	☑	☑
Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	☑	☒	☑	☑
HoNOSCA	☑	☑	☑	☑
Baby Alarm Distress Scale (BADS)	☑	☑	☑	☑
Parent-Infant Relationship Global Assessment Scale (PIR-GAS)	☑	☑ C-GAS and GAF	☑	☑
Parenting-Stress Index – Short Form (PSI/SF)	☑	☒	☑	☑
Strengths and Difficulties Questionnaire	☑	☑ Age	☑	☑

Further details on each of the psychometric tools assessed can be found in Appendix 1.

Discussion

Of the instruments available, the ASQ:SE, PEDS, HoNOSCA, PIR-GAS, PSI/SF and SDQ require further investigation. Although the PEDS appears to be a possibly useful tool, there are limitations. It has been evaluated as possibly cumbersome in an office setting, due to the recommendation for a second screen. Additionally, false positives were found to be higher in some sources (Schonwald, 2007) than others (The Royal Children's Hospital Melbourne, 2012). It does not assess the emotional domain and therefore overall, it appears less suitable.

The BITSEA is designed for 12-36 month old infants and is used as a screening rather than a routine outcome tool. It lacks continuity with the older age group. Additionally, there are copyright issues that would need to be addressed to ensure ready availability and the cumulative cost for routine use would be prohibitive. The ASQ is much broader than the domains of interest while the more applicable ASQ:SE has a complicated protocol for application that potentially interferes with feasibility of national implementation.

The PIR-GAS, SDQ, PSI/SF and BITSEA provide the closest fit with all the criteria. HONOSCA has a role for those aged more than 36 months. The ASQ:SE has a time limitation as it takes 10-15 minutes to complete. However, one if its strengths is its suitability for the 3-47 month infant age range. In The context of NOCC, the ASQ:SE would only need to be used for the 2-36 months age range as the SDQ can address this domain after 36 months. When considering cost, only the HoNOSCA, PIR-GAS and SDQ are free of charge; both the ASQ:SE and BITSEA incur cost. The HoNOSCA and SDQ also provide continuity. Consideration should be given to assessing the HoNOSCA for infants less than 3 years of age. Bearing the strengths and limitations of the available measures in mind, the cost considerations, and balance of criteria, the psychometric measures identified warrant further investigation for potential use as mental health outcome measure in infants 0-47 months in age. The BADS (2-24 months) requires further assessment.

Of the psychometric tools available to date, none cover all of the criteria (Australian Mental Health Outcomes and Classification Network, 2012). Until further investigation is conducted, if instruments were required from the existing group, a combination of measures could be considerd: PIRGAS and PSI/SF across the age group 0-36 months and the ASQ:SE for 3-66 month old infants to measure the social, emotional and behavioural domains of functioning. Where instruments such as HoNOSCA only cover part of the age range, but have a structure that could be applicable to infants from the perspective of clinicians, it would be worth further exploring the development of an infant-appropriate glossary within the clinician based structure of HoNOSCA.

References

Abidin, R. A. (2012). Parenting Stress Index: Short Form (PSI/SF) Retrieved 1 April 2012. [Available from: http://www4.parinc.com/Products/Product.aspx?ProductID=PSI-SF#Items].

Achenbach, T. (2012). Achenbach System of Empirically Based Assessment. Preschool (Ages 11/2-5) Assessments Retrieved 25 March 2012. [Available from: http://www.aseba.org/preschool.html].

Achenbach, T. M., Edelbrock, C. and Howell, C. T. (1987). Empirically based assessment of the behavioral/emotional problems of 2- and 3- year-old children. *Journal of Abnormal Child Psychology,* 15(4)*,* 629-650.

American Psychological Association (2012). Parenting Stress Index Retrieved 1 April 2012, 2012. [Available from: http://www.apa.org/pi/about/publications/caregivers/practice-settings/assessment/tools/parenting-stress.aspx].

Anme, T., Shinohara, R., Sugisawa, Y., Tong, L., Tanaka, E., Watanabe, T., Onda, Y., Kawashima, Y., Hirano, M., Tomisaki, E., Mochizuki, Y., Morita, K., Gan-Yadam, A., Yato, Y., Yamakawa, N. and Japan Children's Study, G. (2010). Interaction Rating Scale (IRS) as an evidence-based practical index of children's social skills and parenting. *Journal of Epidemiology,* 20 Suppl 2*,* S419-426.

Aoki, Y., Zeanah, C. H., Heller, S. S. and Bakshi, S. (2002). Parent-infant relationship global assessment scale: a study of its predictive validity. *Psychiatry & Clinical Neurosciences,* 56(5)*,* 493-497.

Assumpcao, F. B., Jr., Kuczynski, E., Rego, M. G. d. S. and Rocca, C. C. d. A. (2002). [The baby alarm distress scale: a validity study]. *Arquivos de Neuro-Psiquiatria,* 60(1)*,* 56-60.

Australian Mental Health Outcomes and Classification Network (2012). Expert Reference Group Infant Measure Teleconference.

Australian Mental Health Outcomes and Classification Network (2016). Australian Mental Health Outcomes and Classification Network. [Available from: http://www.amhocn.org/background/about-amhocn]

Bayley, N. (2011). Bayley Scales of Infant and Toddler Development Third Edition (BAYLEY-III) Retrieved 25 March 2012. [Available from: http://www.pearsonpsychcorp.com.au/productdetails/34/1/8].

Brann, P., Coleman, G. and Luk, E. (2001). Routine outcome measurement in a child and adolescent mental health service: an evaluation of HoNOSCA. The Health of the Nation Outcome Scales for Children and Adolescents. *Australian & New Zealand Journal of Psychiatry,* 35(3)*,* 370-376.

Briggs-Gowan, M. J. and Carter, A. S. (2007). Applying the Infant-Toddler Social & Emotional Assessment (ITSEA) and BRIEF-ITSEA in Early Intervention. *Infant Mental Health Journal,* 28(6)*,* 564-583.

Briggs-Gowan, M. J. and Carter, A. S. (2008). Social-emotional screening status in early childhood predicts elementary school outcomes. *Pediatrics,* 121(5)*,* 957-962.

Briggs-Gowan, M. J. and Carter, A. S. (2012). Brief comprehensive screening instrument to evaluate social and emotional behavior. Retrieved 25 March, 2012. [Available from: http://www.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=015-8007-352].

Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K. and Cicchetti, D. V. (2004). The Brief Infant-Toddler Social and Emotional Assessment: screening for social-emotional problems and delays in competence. *Journal of Pediatric Psychology,* 29(2)*,* 143-155.

Briggs-Gowan, M. J., Carter, A. S., Skuban, E. M. and Horwitz, S. M. Prevalence of social-emotional and behavioral problems in a community sample of 1- and 2-year-old children. *Journal of the American Academy of Child & Adolescent Psychiatry,* 40(7)*,* 811-819.

Brookes Publishing Company (2012). Ages and Stages Questionnaire (ASQ) Retrieved 13 May 2012. [Available from: http://www.brookespublishing.com/store/books/squires-asq/system.htm].

Brookes Publishing Company (2012a). Ages and Stages Questionnaire. FAQ. Retrieved 1 April 2012. [Available from: http://agesandstages.com/what-is-asq/faq/].

Brookes Publishing Company (2012b). Ages and Stages Questionnaires Third Edition. Retrieved 26 March 2012. [Available from: http://agesandstages.com/what-is-asq/].

Brookes Publishing Company (2012c). Screening and Assessment. [Available from: http://www.brookespublishing.com/tools/asqse/index.htm].

Browne, S., Doran, M. and McGauran, S. (2014). Health of the Nation Outcome Scales (HoNOS): use in an Irish psychiatric outpatient population. *Irish journal of Psychological Medicine, 17(1), 17-19.*

Burgess, P., Pirkis, J. and Coombs, T. (2015). Routine outcome measurement in Australia. *International Review of Psychiatry, 27(4), 264-275.*

Carter, A. and Briggs-Gowan, M. (2012). Infant Toddler Social Emotional Assessment (ITSEA). An in-depth analysis of emerging social-emotional development and intervention guidance. Retrieved 25 March 2012. [Available from: http://www.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=015-8007-387].

Carter, A. S., Briggs-Gowan, M. J. and Davis, N. O. (2004). Assessment of young children's social-emotional development and psychopathology: recent advances and recommendations for practice. *Journal of Child Psychology & Psychiatry & Allied Disciplines,* 45(1)*,* 109-134.

Carter, A. S., Briggs-Gowan, M. J., Jones, S. M. and Little, T. D. (2003). The Infant-Toddler Social and Emotional Assessment (ITSEA): factor structure, reliability, and validity. *Journal of Abnormal Child Psychology,* 31(5)*,* 495-514.

Carter, A. S., Garrity-Rokous, F. E., Chazan-Cohen, R., Little, C. and Briggs-Gowan, M. J. (2001). Maternal depression and comorbidity: predicting early parenting, attachment security, and toddler social-emotional problems and competencies. *Journal of the American Academy of Child & Adolescent Psychiatry,* 40(1)*,* 18-26.

Child and Adolescent Mental Health Outcomes Expert Advisory Panel (2008). Infant Outcome Measures Review. Report to the Child and Adolescent Mental Health Outcomes Expert Advisory Panel Ah Hoc Working Party.

DeGangi, G., Poisson, S., Suckel, R. and Santman, A. (1995). Infant Toddler Symptom Checklist - A Screening Tool for Parents Retrieved 25 March 2012. [Available from: http://www.psychcorp.com.au/productdetails/323].

Diaz-Herrero, A., Lopez-Pina, J. A., Perez-Lopez, J., Brito de la Nuez, A. G. and Martinez-Fuentes, M. T. (2011). Validity of the Parenting Stress Index-Short Form in a sample of Spanish fathers. *Spanish Journal of Psychology,* 14(2)*,* 990-997.

Dydacomp (2005). Achecnbach System of Empirically Based Assessment Retrieved 1 April 2012. [Available from: http://shop1.mailordercentral.com/aseba/prodinfo.asp?number=601].

Eggleston, M. J. F. and Watkins, W. G. A. (2008). Mental health services for children and adolescents in New Zealand, outcomes, and the Health of the Nation Outcome Scale for Children and Adolescents (HoNOSCA). *New Zealand Medical Journal,* 121(1271)*,* 83-91.

Ghuman, J. K., Freund, L., Reiss, A., Serwint, J. and Folstein, S. (1998). Early detection of social interaction problems: development of a social interaction instrument in young children. *Journal of Developmental & Behavioral Pediatrics,* 19(6)*,* 411-419.

Girling, A. (2006). The benefits of using the Neonatal Behavioural Assessment Scale in health visiting practice. *Community Practitioner,* 79(4)*,* 118-120.

Gowers, S. G., Bailey-Rogers, S. J., Shore, A. and Levione, W. (2000). The Health of the Nation Outcome Scales for Child and Adolescent Mental Health (HoNOSCA). *Child Psychology and Psychiatry Review,* 5(2)*,* 50-56.

Gowers, S. G., Harrington, R. C., Whitton, A., Beevor, A., Lelliott, P., Jezzard, R. and Wing, J. K. (1999). Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA). *The British Journal of Psyhiatry,* 174(5)*,* 428-431.

Guedeney, A. and Fermanian, J. (2001). A validity and Reliability study of assessment and screening for sustained withdrawal reaction in infancy: The Alarm Distress Baby Scale. *Infant Mental Health Journal,* 22(5)*,* 559-575.

Haapsamo, H., Ebeling, H., Soini, H., Joskitt, L., Larinen, K., Penninkilampi-Kerola, V., Carter, A. and Moilanen, I. (2009). Screening infants with social and emotional problems: a pilot study of the brief infant toddler social and emotional assessment (BITSEA) in Northern Finland. *International Journal of Circumpolar Health,* 68(4)*,* 386-393.

Hanssen-Bauer, K., Aalen, O. O., Ruud, T. and Heyerdahl, S. (2007). Inter-rater reliability of clinician-rated outcome measures in child and adolescent mental health services. *Administration & Policy in Mental Health,* 34(6)*,* 504-512.

Hanssen-Bauer, K., Gowers, S., Aalen, O. O., Bilenberg, N., Brann, P., Garralda, E., Merry, S. and Heyerdahl, S. (2007). Cross-national reliability of clinician-rated outcome measures in child and adolescent mental health services. *Administration & Policy in Mental Health,* 34(6)*,* 513-518.

Holtmann, M., Becker, A., Banaschewski, T., Rothenberger, A. and Roessner, V. (2011). Psychometric validity of the strengths and difficulties questionnaire-dysregulation profile. *Psychopathology,* 44(1)*,* 53-59.

Karabekiroglu, K., Briggs-Gowan, M. J., Carter, A. S., Rodopman-Arman, A. and Akbas, S. (2010). The clinical validity and reliability of the Brief Infant-Toddler Social and Emotional Assessment (BITSEA). *Infant Behavior & Development,* 33(4)*,* 503-509.

Kerstjens, J. M., Bos, A. F., ten Vergert, E. M. J., de Meer, G., Butcher, P. R. and Reijneveld, S. A. (2009). Support for the global feasibility of the Ages and Stages Questionnaire as developmental screener. *Early Human Development,* 85(7)*,* 443-447.

Kisely, S., Campbell, L. A., Crossman, D., Gleich, S. and Campbell, J. (2007). Are the Health of the Nation Outcome Scales a valid and practical instrument to measure outcomes in North America? A three-site evaluation across Nova Scotia. *Community Mental Health Journal,* 43(2)*,* 91-107

Kowalenko, N. and Reilly, N. (2007). Working Papers: Discussion Points and Comments for Teleconference.

Kruizinga, I., Jansen, W., Carter, A. S. and Raat, H. (2011). Evaluation of an early detection tool for social-emotional and behavioral problems in toddlers: The Brief Infant Toddler Social and Emotional Assessment - A cluster randomized trial. *BMC Public Health,* 11*,* 494.

Leone, S. L. (2009). *Evaluation of the Ghuman-Folstein Screen for Social Interaction (SSI).* The Ohio State University.

Lesinskiene, S., Senina, J. and Ranceva, N. (2007). Use of the HoNOSCA scale in the teamwork of inpatient child psychiatry unit. *Journal of Psychiatric & Mental Health Nursing,* 14(8)*,* 727-733.

Limbos, M. M. and Joyce, D. P. (2011). Comparison of the ASQ and PEDS in screening for developmental delay in children presenting for primary care. *Journal of Developmental & Behavioral Pediatrics,* 32(7)*,* 499-511.

Mackrides, P. S. and Ryherd, S. J. (2011). Screening for developmental delay. *American Family Physician,* 84(5)*,* 544-549.

Mantymaa, M., Puura, K., Luoma, I., Kaukonen, P., Salmelin, R. K. and Tamminen, T. (2008). Infants' social withdrawal and parents' mental health. *Infant Behavior & Development,* 31(4)*,* 606-613.

Marks, K. (2007). Should general pediatricians not select the Ages & Stages Questionnaire in light of the Rydz et al study? *Pediatrics,* 120(2)*,* 457-458; author reply 458; discussion 459.

Monroe Carell Jr. Children's Hospital at Vanderbilt (2012). Appointments Retrieved 1 April 2012. [Available from: http://www.childrenshospital.vanderbilt.org/interior.php?mid=344].

Schonwald, A. (2007). Developmental Screening Tool Kit for Primary Care Providers Retrieved 26 March 2012. [Available from: http://www.developmentalscreening.org/screening_tools/peds.htm].

Skovgaard, A., Olsen, E., Christiansen, E., Houmann, T., Landorph, S. and Jorgensen, T. (2008). Predictors (0-10 months) of psychopathology at age 1 1/2 years--A general population study in the Copenhagen Child Cohort CCC 2000*. *Journal of Child Psychology and Psychiatry,* 49(5)*,* 553-562.

Smith, M. K. and Brun, C. F. (2006). An analysis of selected measures of child well-being for use at school-and community-based family resource centers. *Child Welfare,* 85(6)*,* 985-1010.

The California Evidence-Based Clearinghouse for Child Welfare (2012, June 2009). Screening and Assessment Tools for Child Welfare. Pediatric Symptom Checklist-17 (PSC-17) Retrieved 19 November 2012, 2012. [Available from: from http://www.cebc4cw.org/assessment-tool/pediatric-symptom-checklist-17/].

The Royal Children's Hospital Melbourne (2012). Child Development Assessment tool. Parents' Evaluation of Developmental Status (PEDS) Retrieved 6 May 2019. [Available from: http://www.rch.org.au/ccch/resources.cfm?doc_id=10963].

The Washington Institute and the Washington State Mental Health Division (2012). Parent-Infant Relationship Global Assessment (PIR-GAS). [Available from: http://depts.washington.edu/washinst/Resources/CGAS/PIRGAS%20Index.htm].

UMDNJ, V. I. o. N. J. a. (2006). Searchable Inventory of Instruments. Assessing Violent Behavior and Related Constructs in Children and Adolescents Retrieved 1 April 2012. [Available from: http://vinst.umdnj.edu/VAID/TestReport.asp?Code=CTRF].

youthinmind (2012). Youth In Mind Retrieved 1 April 2012. [Available from: http://www.sdqinfo.com/py/doc/f0.py].

Appendix 1: Summary of Psychometric Tools Assessed against Evaluation Criteria

This table is a summary of the assessment of the listed psychometric tools against whether each of the evaluation criteria were met; Yes = ☑ No = ☒ Potential = ? Source: webpages, medline and psychinfo; as per reference list below.

Psycho-metric tool	**Age** 0-47 months	**Comple-tion Rate** ≤50 items and <10 mins	**Social/ Emotional/ Behavioral** Explicitly measures changes on infants 0-47 months.	**Avail-able** Cost, copyright	**Sound** Internal consistency, validity, reliability and sensitivity to change	**Scient-ific** Developed, piloted, scrutinised	**Continuity** With measures used in later life	**Quantitative**	**Generalisable** To Australian context
Ages and Stages Questionnaire: Social-Emotional (ASQ:SE)	☑ 0-5 ½ years (Brookes Publishing Company, 2012c; Mackrides and Ryherd, 2011) Used in study for children 12-60 months(Limbos and Joyce, 2011, p. 501) Age range: 1–66 months for ASQ-3, 3–66 months	? 30 questions, 15 minutes to complete (Limbos and Joyce, 2011, p. 501) 10-15 mins (Mackrides and Ryherd, 2011) Number of questionnaires: 21 for ASQ-3, 8 for ASQ:SE (Brookes Publishing Company, 2012a)	☑ Social/emotional, developmental delays (Brookes Publishing Company, 2012c)	☑ Low-cost (Brookes Publishing Company, 2012d)	☑ Good psychometric properties (Kerstjens, Bos, ten Vergert, et al., 2009) Described as one of two "…most extensively evaluated parent-completed tools…" (Mackrides and Ryherd, 2011)	☑ Scientifically scrutinised (Marks, 2007)	☒	☑ Uses scoresheets (Brookes Publishing Company, 2012c)	☑ Useful worldwide (Kerstjens, Bos, ten Vergert, et al., 2009)

	for ASQ:SE (Brookes Publishing Company, 2012a)				Test-retest reliability = 0.92, Inter-rater reliability=0.93. Validity=0.82-0.88, Sensitivity = 0.86, Specificity=0.85 (Brookes Publishing Company, 2012b).				

Psychometric tool	Age	Completion Rate	Social/ Emotional/ Behavioral	Available	Sound	Scientific	Continuity	Quantitative	Generalisable
Parents' Evaluation of Developmental Stages (PEDS)	☑ 0-8 years	☑ 10 items, few mins (The Royal Children's Hospital Melbourne, 2012). Second screen for children with only one significant concern suggests as possibly cumbersome ion office settings (Schonwald, 2007)	? For developmental and behavioural problems (The Royal Children's Hospital Melbourne, 2012) Designed to identify children eligible for special education (Schonwald, 2007)	☑ $55 for set of 50.(The Royal Children's Hospital Melbourne, 2012)	☑ Sensitivity and specificity 70-80% (The Royal Children's Hospital Melbourne, 2012) False positives are high (Schonwald, 2007)	☑ Subjected to Scientific scrutiny (Limbos and Joyce, 2011)	☒	☑ Score form (The Royal Children's Hospital Melbourne, 2012)	☑ Yes (The Royal Children's Hospital Melbourne, 2012)

Psychometric tool	Age	Completion Rate	Social/ Emotional/ Behavioral	Available	Sound	Scientific	Continuity	Quantitative	Generalisable
ITSEA	? 12-36 months (A. Carter and Briggs-Gowan, 2012)	☒ 25-30 mins							
BITSEA	? 12-35 months (M.J. Briggs-Gowan and Carter, 2012)	☑ 7-10 mins, 42 items (Turkish version) (M.J. Briggs-Gowan and Carter, 2012)	☑ Externalising, Internalising, Dysregulation, Competence (M.J. Briggs-Gowan and Carter, 2012)	☑ $110.25 for set of 25 forms	☑ Test-retest reliability excellent, high internal consistency (Cronbach's α = 0.80 for problem scale, Cronbach's α=0.69 for competence scale). Inter-rater reliability statistically significant. Found to be a valid and reliable measure (Karabekiroglu, Briggs-Gowan, Carter, et al.,	☑ Undergone scientific scrutiny (Margaret J. Briggs-Gowan, Carter, Irwin, et al., 2004)	☒	☑ Yes. (M.J. Briggs-Gowan and Carter, 2007, p. 571)	☑ Yes.

					2010, pp. 504, 506-508) .				

Psychometric tool	Age	Completion Rate	Social/ Emotional/ Behavioral	Available	Sound	Scientific	Continuity	Quantitative	Generalisable
Ghuman-Folstein Screen for Social Interaction	? 2-5 years	☒ 54 items	☒ Screen for social interaction for diagnosing autism spectrum disorders(Leone, 2009)						

HoNOSCA	? 3-18 years	☑ 8 mins (range: 2-20 mins)(Gowers, Bailey-Rogers, Shore, et al., 2000)	☑ Psycho-social, emotional (Hanssen-Bauer, Aalen, Ruud, et al., 2007)	☑ Cost-effective (Eggleston and Watkins, 2008)	☑ Good psychometric properties. Interclass correlation co-efficient = 0.84 (Hanssen-Bauer, Gowers, Aalen, et al., 2007). Verified as satisfactory elsewhere (Lesinskiene, Senina and Ranceva, 2007).	☑ Scrutinised under peer review (Eggleston and Watkins, 2008; S.G. Gowers, R.C. Harrington, A. Whitton, et al., 1999; Hanssen-Bauer, Aalen, Ruud, et al., 2007; Hanssen-Bauer, Gowers, Aalen, et al., 2007; Kisely, Campbell, Crossman, et al., 2007; Lesinskiene, Senina and Ranceva, 2007)	☑	☑ Questions yield quantitative results (S.G. Gowers, R.C. Harrington, A. Whitton, et al., 1999)	☑ Yes; English. Also used in New Zealand.

Neonatal Behavioral Assessment Scale (NBAS)	☒ 0-2 months (Girling, 2006, p. 119)	N/A	☒ Not relevant – for neonates to test for sleep, reflex, auditory and visual stimuli.						
Denver Developmental Screening Test (DDST)	☑ 2 weeks to 6 years (Schonwald, 2007)	☒ 20 minutes required (Schonwald, 2007)							

Psychometric tool	Age	Completion Rate	Social/ Emotional/ Behavioral	Available	Sound	Scientific	Continuity	Quantitative	Generalisable
Child Behavior Checklist	? 1.5-5 years 1 ½ to 5 years (part of ASEBA)	☒ 99 items reported. (Alice S. Carter, Briggs-Gowan and Davis, 2004) >50 questions total counted on a copy of questionnaire. Forty-three questions are likert scale type (Monroe Carell Jr. Children's Hospital at Vanderbilt, 2012)	☑ Social, emotional, behavioral, communication and motor skills (Monroe Carell Jr. Children's Hospital at Vanderbilt, 2012)	☑ $25.00 for pack of 50 (Dydacomp, 2005)	☑ Good test-retest reliability (Alice S. Carter, Briggs-Gowan and Davis, 2004)	☑ Peer-reviewed (Margaret J. Briggs-Gowan and Carter, 2008)		☑ t-scores and cut-point scores given (Alice S. Carter, Briggs-Gowan and Davis, 2004)	☑ Yes.

Psychometric tool	Age	Completion Rate	Social/ Emotional/ Behavioral	Available	Sound	Scientific	Continuity	Quantitative	Generalisable
Infant and Toddler Symptom Checklist	☒ 7-30 months	☑ 10 minutes	☑ Screen for regulatory and sensory disorders; self-regulation, attention, modulation of sleep/wake status, responses to sensory stimulation, attachment, and emotional functioning (DeGangi, Poisson, Suckel, et al., 1995)	☑ $216 for complete kit					

Psychometric tool	Age	Completion Rate	Social/ Emotional/ Behavioral	Available	Sound	Scientific	Continuity	Quantitative	Generalisable
Bayley Scales of Infant and Toddler Development (BSID III)	☑ 1-42 months (Bayley, 2011)	☒ 30-60 mins	☑ Cognition, language, social-emotional, motor, adaptive behavior.	☒ Cost is high. $2033 for complete kit with 25 forms.					

Psychometric tool	Age	Completion Rate	Social/ Emotional/ Behavioral	Available	Sound	Scientific	Continuity	Quantitative	Generalisable
Baby Alarm Distress Scale	?	☑	☑	☑	☑	☑	☑	☑	☑
	2-24 months only	14 items	Facial expression, eye contact, general activity, self-stimulating gestures, vocalizations, response to stimulation, relationship to observer, ability to attract attention (Guedeney and Fermanian, 2001, p. 563)	Free	Good validity, sensitivity = 0.82 and specificity =0.78. Good construct validity, satisfactory reliability.	Scientifically scrutinized (Guedeney and Fermanian, 2001, p. 564)and used in other studies (Assumpcao, Kuczynski, Rego, et al., 2002)		Yes.	Yes. Translated to English from the original French, and blind retranslated into French.

Psychometric tool	Age	Completion Rate	Social/ Emotional/ Behavioral	Available	Sound	Scientific	Continuity	Quantitative	Generalisable
Interaction Rating Scale	? Participants in study were 18, 30, 42 months. Unclear whether can be used I <18 months	☒ Too lengthy. 70 items for behavioural score, plus 11 items for impression score.	☑ Social competence		☑ Inter-rater reliability = 90% (Anme, Shinohara, Sugisawa, et al., 2010, pp. S-420). Internal consistency.				

Psychometric tool	Age	Completion Rate	Social/ Emotional/ Behavioral	Available	Sound	Scientific	Continuity	Quantitative	Generalisable
Caregiver-Teacher Report Form (C-TRF 2-5) under the Achenbach System of Empirically Based Assessment (ASEBA)	? 2 to 5 years	☒ 99 items (UMDNJ, 2006)	☑ Social, emotional, behavioral	☑ 22 GBP for pack of 50	☑ Test-retest reliability 0.64 to 0.91, Internal consistency 0.52 to 0.9. Criterion validity acceptable (UMDNJ, 2006)				

Psychometric tool	Age	Completion Rate	Social/ Emotional/ Behavioral	Available	Sound	Scientific	Continuity	Quantitative	Generalisable
Parent-Infant Relationship Global Assessment Scale (PIR-GAS)	☑ 0-5 years	☑ 90-point scale in 10 parts (The Washington Institute and the Washington State Mental Health Division, 2012)	☑ Parent-infant relationship (The Washington Institute and the Washington State Mental Health Division, 2012)	? Cost is unclear	☑ Predictive validity demonstrated (Aoki, Zeanah, Heller, et al., 2002)	☑ Scientifically scrutinised (Skovgaard, Olsen, Christiansen, et al., 2008)	☑	☑ Quantitative scale	☑ Yes.
Parenting-Stress Index – Short Form (PSI/SF)	☑ 0-3 years	☑ Short form (36 items; 10 mins) has 3 subscales; parental distress, parent-child dysfunctional interaction,	☑ Child subscales: Adaptability, Acceptability, Distractability/Hyperactivity, Demandingness, Mood, Reinforces Parent Parent subscales: Competence, Social Isolation,	☑ $126.00 for kit with 25 forms (Abidin, 2012).	☑ Very good reliability parent =0.55-0.80; child 0.62-0.70. Test-retest reliability after 1 year: 0.70 Parent, 0.55 Child; After 3 weeks: 0.71	☑ Undergone peer review (Diaz-Herrero, Lopez-Pina, Perez-Lopez, et al., 2011)		☑ Total Stress Score	☑ Available in English

		difficult child (American Psychological Association, 2012).	Attachment, Parent Health, Role Restriction, Depression, Relationship with Spouse (American Psychological Association, 2012)		Parent and 0.82 Child. Validity: 41% on child section accounted for by 6 factors, 44% on parent section accounted for by 7 factors (American Psychological Association, 2012)				
Strengths and Difficulties Questionnaire	? 3-16 years	☑ 33 questions	☑ Emotional, conduct, hyperactivity/inattention, peer relationship problems, prosocial behavior (youthinmind, 2012)	☑ Free of charge for non-commercial purposes (youthinmind, 2012)	☑ Psychometric properties robustly tested and validated (Holtmann, Becker, Banaschewski et al., 2011)	☑ Peer-reviewed(Holtmann, Becker, Banaschewski, et al., 2011)	☑	☑ Yes.	☑ Has Australian version (youthinmind, 2012)
Pediatric Symptom Checklist-17	☒ 4-18 years								

End of Report.