

The Longitudinal Study of Australian Children: An Australian Government Initiative

LSAC Technical Paper #2

Summarising children's wellbeing: the LSAC Outcome Index

Ann Sanson, Sebastian Misson and other members of the Outcome Index Working Group

(Melissa Wake, Stephen R Zubrick, Sven Silburn, Sheldon Rothman, Jane Dickenson)

September 2005



Australian Government

Australian Institute of Family Studies

Table of Contents

Table of Contents	2
About the authors	3
Introduction	4
Framework for Conceptualising Child Outcomes	5
Approach to measurement	7
Selection of component variables	8
Calculation of the Outcome Index	10
Illustrative examples of the Use of the Index	17
Conclusions	22
References	23
Appendix A: Handling missing data	24
Appendix B: Summary of variables included in the Outcome	
Index	27

About the authors

Associate Professor Ann Sanson, LSAC Project Director, University of Melbourne. Ann was at the Australian Institute of Family Studies from 2000 to 2004, on secondment from the University of Melbourne. She held the roles of Principal Research Fellow, Deputy Director and Acting Director during this time. As LSAC Project Director, she also led the early development work on LSAC. She now fills the role of Project Director on a part-time basis.

Mr Sebastian Misson is employed at the Australian Institute of Family Studies as Data Administrator for LSAC.

Dr Melissa Wake is Director of the Research and Public Health Unit at the Centre for Community Child Health, Murdoch Childrens Research Institute.

Professor Stephen R Zubrick is Head of the Division of Population Science at the Telethon Institute for Child Health Research in Western Australia, and Co-Director of the Centre for Developmental Health at Curtin University of Technology.

Professor Sven Silburn is Director of the Centre for Developmental Health at Curtin University and a Senior Research Fellow in the Division of Population Science at the Telethon Institute for Child Health Research, Western Australia.

Dr Sheldon Rothman is a Principal Research Fellow at the Australian Council for Educational Research, where he manages the Longitudinal Surveys of Australian Youth project.

Jane Dickenson is a member of the LSAC team at the Department of Family and Community Services

Introduction

This paper describes the development of the Outcome Index for Wave 1 of *Growing Up in Australia* - the Longitudinal Study of Australian Children (LSAC), and illustrates some of the ways in which it can be used to shed light on how well children are faring and how this relates to the conditions of their life. *Growing Up in Australia* is funded by the Australian Government Department of Family and Community Services (FaCS) as part of the Australian Government's *Stronger Families and Communities Strategy*. The Australian Institute of Family Studies is leading a consortium of nine Australian research institutions in the development of this study, which will track the development of two cohorts of young children for at least 7 years.

Growing Up in Australia is one of the largest and most complex studies of this nature that has ever been undertaken in Australia. The study aims to provide the data for a comprehensive understanding of Australian children's development in the current and future social, economic and cultural environment, and hence to become a major element of the evidence base for policy and practice regarding children and their families.

The Outcome Index is a composite measure to indicate how children are developing. LSAC tracks the development of children across multiple domains, and the Outcome Index provides a means of summarising this complex information for policy makers, the media and the general public, as well as potential data users. In contrast to other indices which focus on problems, such as the Vulnerability Index developed in the Canadian National Longitudinal Study of Children and Youth (NLSCY; Willms, 2002), the LSAC Outcome Index wherever possible incorporates both strengths and weaknesses, reflecting the fact that most children have good developmental outcomes. Thus the Outcome Index has the ability to identify groups of children developing poorly and those developing well.

In developing the LSAC Outcome Index, the following guiding principles were followed:

- 1. The Index should provide a snapshot of the children's development at the time of each main wave of data collection.
- 2. The Index should contain all dimensions of developmental interest.
- 3. The Index should have policy and research relevance.
- 4. The Index should be made up of constructs that can be measured throughout childhood and into adulthood.
- 5. The Index should be able to identify both those children doing poorly and those doing well, wherever possible.
- 6. The Index should be predictive of later childhood, adolescent and adult outcomes.
- 7. The Index should be parsimonious.
- 8. The Index should not include distal factors (e.g. income, family structure) that are predictively related to child outcomes, but rather focus on actual developmental status.

The Outcome Index Working Group has undertaken the development work, led by Ann Sanson and Sebastian Misson. Besides having responsibility to ensure the validity of the Index, the Working Group has tried also to take the perspective of end-users of the Index, both as researchers and as policy-makers.

This paper first presents the conceptual framework underlying the Index, and the broad approach to measurement adopted. The specific variables incorporated in the Index are then discussed, followed by a description of its derivation. Some illustrative examples of the use of the Index are then presented.

Framework for Conceptualising Child Outcomes

For the purpose of the LSAC Outcome Index, an outcome is an attribute of the child at a particular point in time. Hence, factors that may influence child development and are sometimes used as indicators of child wellbeing (e.g. poverty, maternal depression, prematurity) are not included as they are not direct measures of child development. If such indicators were to form part of the Index, it would not be possible to test empirically their influence on child wellbeing, It is therefore important to make a clear separation between actual current child functioning and possible causal influences.

In developing the LSAC Outcome Index we consider children both as "beings" and as "becomings". As "beings", we are concerned about children's wellbeing here and now, at a particular point in time – the domains of social, emotional, physical/health and school-related functioning capture the major aspects of wellbeing of a child. As "becomings", we need to consider the eventual outcomes we would like children to develop towards. It is proposed that a key outcome of interest is that children become adults capable of productive social and economic participation in society. This encompasses their future roles as learners, workers, family members and community members. While there are many other outcomes we might hope children will achieve (e.g. happiness), the broad goals of economic and social participation may best capture what it means to be a positive contributor to society.

The broad framework for the LSAC Outcome Index can be seen in Figure 1. Three domains are proposed to be the major components of current wellbeing and the future capability to be a successful civic and economic participant: health and physical development, social and emotional functioning, and learning and academic competency. Summary scores for each of these domains are calculated, and they are combined into the overall Index. Since each of these domains is roughly equally important for a child's developmental wellbeing, they are equally weighted in the overall Index. It is acknowledged that children's development is multidimensional and interactive, and that there are "fuzzy boundaries" between these domains of functioning. At the same time, a child's development may not be uniform across domains; for example, a child may be doing well in language, but have poorly developed motor skills. It is important to capture this variability.

A criterion for components in the Index is that they should be measurable from Wave 1 (for at least the four-year-old cohort) to adulthood. Measures of child functioning obviously differ by age, so the specific measures of many components will change over waves, but their meaning, and the meaning of the overall Index, will be broadly consistent across the age span.

Figure 1: Conceptual framework for Outcome Index for 0 and 4-year-olds, showing domains (in uppercase) and subdomains (in lower case)



A second criterion is that the Index should be predictive of later outcomes. While the variables selected are those that the literature suggests will be predictive, it is obviously not possible to test this criterion empirically at this stage. It is possible that later waves of data will suggest that some revision of the Index would be worthwhile in order to increase its predictiveness.

As can be seen from Figure 1, not all of the components that can be measured for the 4year-old sample can be measured in the infant sample, since these outcomes are not observable at such a young age, and/or are not able to be assessed in LSAC. Hence the meaning of the Index varies to some extent across cohorts, at least in Wave 1. This needs to be taken into consideration when making use of the infant Index, particularly in conjunction with Indices from later waves.

Approach to measurement

In researching the most appropriate way to develop the Outcome Index, two approaches to combining variables for each subdomain and domain were initially investigated.

The first was modelled on the NLSCY Vulnerability Index and involved the creation of cut-off scores for each variable to indicate problem/no-problem status. An iterative process involved identifying the number of *variables* within a subdomain where the child had problem status, and then (through the use of another cut-off) the number of *subdomains* where the child had problem status, and finally the number of *domains* where the child had problem status. These last could be summed to create the 'problem' index. Since it was desired to also identify children who were doing particularly well, the same process would be used for creating a positive index.

The major limitation of this approach is that it requires a large number of essentially arbitrary decisions about where cut-offs should be placed at each stage of the process. Further, the nature of the distribution of scores across variables led to different proportions of the sample being identified as having problem (or positive) status across variables and domains, thus resulting in uneven weighting of variables and domains within the Index. Finally, the description of the derivation of the Index would be complex and may deter potential users who found it difficult to understand.

The second approach retains variables in their continuous form, but uses standardisation of all contributing variables to make them comparable. A subdomain score is computed as the sum of standardised scores on variables contributing to it, and a domain score is the sum of standardised subdomain scores. A continuous Outcome Index score could be derived from standardisation of these domain scores. A categorical form of the OI could also be computed by applying cut-offs (e,g. top and bottom 15%) to the domain scores to identify positive and problem status and then summing the number of domains with problem status, and similarly summing the number with positive status. The process is described in detail in the "Calculation of the Outcome Index" section below.

The major advantage of this continuous approach is that there is no reliance on cut-offs until the last stage of computation, and hence information is not lost through reducing continuous scores on variables to dichotomous categories. Further, each sub-domain within a domain, and each domain within the overall Outcome Index, can be approximately evenly weighted so that each contributes equally, and explanation of how the Index is calculated is relatively straightforward. Given these advantages, this approach was adopted.

Selection of component variables

For ease of interpretability, it was desirable to include a small set of reliable variables in the Outcome Index. Given the scope of the LSAC dataset, there were a considerable number of candidate variables to consider for inclusion. The following criteria were used in choosing among these:

- Variables with high response rates (derived from the Parent 1 interview) were preferred over those with lower response rates (Parent 1 self-complete, teacher or carer data). If it was possible to get good coverage of a domain with high-response variables, only these were used. However, for some content areas only lower response variables were available, and on others it was considered desirable to include two informants (e.g. parent and teacher). Procedures to manage missing data, when necessary, are described in Appendix A.
- Distribution of scores on the variable allowed discrimination between those with poor and average functioning, and preferably also between those with average and good functioning. In general terms, it was assumed that about 5-20% of the sample should be identified as doing poorly on any one variable, and a similar proportion as doing well. Some variables were excluded because they were not sufficiently discriminating (e.g. 50% of 4 year olds had sleep problems, and 40% had injuries; parents of 99% of infants had no concerns about their infants' fine motor coordination). Some selected variables gave good discrimination between poor and average but less adequate discrimination between average and good.
- Adequate internal consistency of scales, where relevant.
- Age trends were also identified (i.e. where younger children in a cohort performed differently to older children in the cohort), so that these could be taken into account in computing the Index.

Item content of the reduced set of candidate variables for each subdomain was then examined along with interrelationships among them (via correlations and cross-tabs) to ensure good conceptual coverage of the subdomain without redundancy.

The final set of variables thus selected for inclusion in the Outcome Index is shown in Figure 2. Full descriptions of these variables can be found in the LSAC Data Users Guide and brief summaries are provided in Appendix B.



Calculation of the Outcome Index

(a) Four-Year-Olds

The calculation of the Outcome Index for the 4-year-olds is a 4 stage process. Stage 1 involves standardising all the outcome variables and combining them into sub-domain scores. The second involves standardising the sub-domain scores and combining them in domain scores. The third stage involves standardising the domain scores, and obtaining cut-offs to identify the top 15% and bottom 15% of the sample for each domain. The final stage calculates the final Outcome Indices by identifying the number of domains on which a child is in either the top 15%, or the bottom 15%, and also calculating a continuous index score by averaging the three domain scores. All analyses and derivations involved use weighted data where appropriate.

Stage 1

Descriptive statistics for the variables included in the Outcome Index for the 4-yearolds can be seen in Table 1. All variables were standardised to have a mean of 0 and a standard deviation of 1 (i.e. z-scores), so that all would be of equal weight. Age effects were observed for the WAI, PPVT and parent rating of reading ability. To ensure that age of assessment did not affect the index, these variables were standardised by age group using 5 categories: less than 54 months, 55-56 months, 57-58 months, 59-60 months and more than 60 months. Those variables where high scores indicated a poorer outcome were then multiplied by -1, so that high scores always indicated better outcomes. In order to reflect concerns about both underweight and overweight children, the absolute value of the z-score of the Body Mass Index was obtained by multiplying the scores of children with a z-score below 0 by -1, so that children who were significantly below the mean and those above the mean both had higher scores that children with average BMI. This score was then standardised and reversed so that, like the other variables, it was a z-score with high scores representing a better outcome (i.e. average weight).

	Ν	Mean	SD
Physical			
Health			
Overall rating of health	4982	1.60	0.79
Special health-care needs screening	4934	1.86	0.34
Body mass index	4934	16.37	2.01
Motor			
PEDS QL Physical health summary	4198	82.60	12.43
Social/Emotional			
Social Competence			
SDQ Prosocial	4969	7.73	1.80
SDQ Peer problems	4969	1.72	1.58
Internalising			
SDQ Emotional symptoms	4968	1.75	1.71
Externalising			
SDQ Hyperactivity	4969	3.59	2.29

Table 1. Descriptive statistics of outcome variables for 4-year-olds

SDQ Conduct	4969	1.10	1.77
Learning			
Language			
ЎРРVТ	4382	63.63	8.23
Literacy			
Parent rating of reading skills	4971	0.53	0.84
Teacher rating of reading skills	3206	2.01	1.09
Teacher rating of writing skills	3216	3.45	1.52
Numeracy			
Teacher rating of numeracy skills	3194	3.59	1.26
Approach to learning			
WAI	4880	63.96	8.11

For sub-domains with only one contributing variable (namely, the Motor subdomain of Physical, the Internalising subdomain of Social/Emotional, and the Language and Numeracy subdomains within Learning), the z-score for that variable became the subdomain score. For sub-domains with more than one contributing measure, the mean of z-scores for contributing variables was then obtained to form a sub-domain score. As is evident in Table 1, there was some missing data for the Motor subdomain, and a considerable amount of missing teacher data (this includes cases where the child did not attend childcare, preschool or school, cases where parents withheld permission to contact teachers/carers, and cases where teachers did not return the questionnaire). How these missing data were handled is described in Appendix A.

Stage 2

Stage 2 involved using the standardised sub-domain scores to create the domain scores, essentially by repeating the process above. Domain scores were created by taking the average of the relevant sub-domain scores. Missing data was again an issue in the creation of domain scores. While only one case had one missing subdomain score in the Social/Emotional domain, a total of 805 cases had one missing subdomain subdomain score in the Physical domain, and a large number did not have teacher data available for some subdomains of Learning (n=2088). Again, the process described in Appendix A was followed to deal with missing data.

Stage 3

Following standardisation by missingness, all three domain scores were standardised again to have a mean of 100 and a standard deviation of 10 for greater ease of use. Cut-offs were then identified for the top and bottom 15% on each domain score. The cut-offs can be seen in Figures 3 to 5, with the cases represented in red on the figures falling in the top or bottom 15%.



Figure 3: Distribution of Physical domain continuous scores for four-year-olds



Figure 4: Distribution of Social/Emotional domain continuous scores for fouryear-olds



Figure 5: Distribution of Learning domain continuous scores for four-year-olds

Stage 4

The overall Outcome Index can be represented in two forms:

- As a continuous score, to which cut-offs can be applied to identify those with poor (bottom 15%) and superior (top 15%) functioning. The continuous Outcome Index was calculated by obtaining the average of the 3 domain scores (in cases where one of these domain scores was missing, no average was obtained, and the child was 'missing' on the Outcome Index). It was then standardised to have a mean of 100 and a standard deviation of 10 for greater ease of use.
- Two categorical variables were created, representing the total number of domains in which a child scores below the negative cut-off (Negative Outcome Index), or above the positive cut-off (Positive Outcome Index). For the positive Outcome Index, 63.9% of children scored above none of the cut-offs, 28.0% above one, 7.4% above two and 0.7% above all three. For the negative Outcome Index, 66.0% scored below none of the cut-offs, 24.6% below one, 7.6% below 2 and 1.7% below all three. Table 2 shows the distribution when the Negative and Positive Indices are considered together. While over a third of children have neither strengths nor weaknesses, as would be expected, just under one quarter have one area of difficulty, and 9.3% have 2-3 areas of difficulty; 28.0% have one area of strength, and 8.2% have 2-3 areas of strength. It can be seen that 5.9% of children show some strengths but also have weaknesses. It is also possible to identify subgroups of children with particular patterns of outcomes across domains; e.g. those with problems in the Physical domain but strengths in the Learning domain, or those with average Physical and Social/Emotional functioning but poor Learning.

			Number of positive domains						
		0	1	2	3	Total			
Number	0	1781	1125	345	35	3286			
of		(35.9%)	(22.7%)	(6.9%)	(0.7%)	(66.2%)			
negative	1	957	235	27		1219			
domains		(19.3%)	(4.7%)	(0.6%)		(24.6%)			
	2	345	32			377			
		(7.0%)	(0.6%)			(7.6%)			
	3	84				84			
		(1.7%)				(1.7%)			
	Total	3168	1391	372	35	4966			
		(63.8%)	(28.0%)	(7.5%)	(0.7%)	(100.0%)			

Table 2: Distribution of 4-year-olds across Positive and Negative Indices

Infants

The calculation of the Outcome Index for the infants is a slightly simpler 3-stage process since there are no sub-domains (see Figure 1). For the infants, stage 1 involves standardising all the outcome variables and combining them into domain scores. The second stage involves standardising the domain scores, and obtaining cut-offs to identify the top 15% and the bottom 15% of the sample for each domain. The final stage calculates the final Outcome Indices by identifying the number of domains in which the child is in either the top or bottom 15%, and also calculating a continuous Index score by averaging the three domain scores. Again, all analyses and derivations involved use weighted data where appropriate.

Stage 1

Descriptive statistics of the outcome variables used for the infants can be seen in Table 3. As for the 4-year-olds, all outcome variables were standardised to have a mean of 0 and a standard deviation of 1 (i.e. z-scores) so that all variables would be of equal weight. There were four variables on which age effects were observed (namely, special health care needs, STSI Approach, STSI Cooperativeness, and the CSBS), and these were standardised by age group to take account of age at assessment. The 5 age groups used were: less than 6 months, 6 to 7 months, 8 to 9 months, 10 to 11 months and 12 months or more. After this standardisation those variables where high scores indicated negative outcomes were multiplied by -1 so that high scores indicated positive outcomes for all variables.

Table 3.	Descriptive	statistics	of outcome	variables t	for infants
----------	-------------	------------	------------	-------------	-------------

	Ν	Mean	SD
Physical			
Overall rating of health	5106	1.55	0.80
Special health-care needs screening	5029	1.94	0.24
Social/Emotional			
STSC Irritability	4311	2.50	0.82
STSC Approach	4314	4.72	0.85
STSC Cooperativeness	4312	4.17	0.89
Learning			

	CSBS standardised score 4445	99.51	15.55
--	------------------------------	-------	-------

Stage 2

The mean of these standardised scores was then obtained for each domain. The Learning domain contained only one, already-standardised variable, hence no further standardisation was necessary. As can be seen from Table 3, there were very few missing data. However, for consistency, the process used for 4-year-olds was adopted to adjust for missingness, as outlined in Appendix A.

All three sub-domain scores were then standardised to have a mean of 100 and a standard deviation of 10 for greater ease of use. Due to the nature of the distributions, it was not always possible to identify exactly 15% of the sample for these cut-offs, but they were placed as close to 15% as possible. The cut-offs for the 3 domains can be seen in Figures 6 to 8, along with the proportion of the sample they identify. Figure 6 shows that no 'positive' cut-off was meaningful.



Figure 6: Distribution of Physical domain continuous scores for infants

Figure 7: Distribution of Social/Emotional domain continuous scores for infants

Figure 8: Distribution of Learning domain continuous scores for infants

Stage 3

Again, the overall Outcome Index can be represented in two forms:

- As a continuous score, to which cut-offs can be applied to identify those with poor (bottom 15%) and superior (top 15%) functioning, as for 4-year olds.
- Two categorical variables, representing the total number of domains in which a child scores below the negative cut-off (Negative Outcome Index), or above the positive cut-off (Positive Outcome Index). As noted above, given that

58.6% of the sample achieved the best possible health outcomes (i.e. 'Excellent' rating of overall health and no special health care needs), it was not possible to develop a reasonable positive cut-off for this domain. Therefore only two domains (Social/Emotional and Learning) are included in the final categorical Positive Outcome Index, with a range of 0 to 2. For the Positive Outcome Index, 73.4% were identified by none of the cut-offs, 24.1% by one of the cut-offs and 2.6% by both of the cut-offs. For the Negative Outcome Index, 63.5% of children were below none of the cut-offs, 28.7% on 1 of the cut-offs, 7.0% on 2 of the cut-offs and 0.8% on all three of the cutoffs. Table 4 shows the distribution when the Negative and Positive Indices are considered together. Over 40% of children have neither strengths nor weaknesses; over one quarter have one area of difficulty, and 7.8% have 2-3 areas of difficulty; just under one quarter have one area of strength, and 2.6% have 2 areas of strength. Similar to the 4-year-olds, 6.3% of children show some strengths as well as weaknesses.

			Number of positive domains						
		0	1	2	Total				
Number	0	1615	669	89	2374				
01		(43.2/0)	(17.9%)	(2.4%)	(63.5%)				
negative	1	856	210	8	1074				
domains		(22.9%)	(5.6%)	(0.2%)	(28.7%)				
	2	243	20		263				
		(6.5%)	(0.5%)		(7.0%)				
	3	29			29				
		(0.8%)			(0.8%)				
	Total	2743	900	97	3740				
		(73.4%)	(24.0%)	(2.6%)	(100%)				

Table 4: Distribution of infants across Positive and Negative Indices

Illustrative examples of the Use of the Index

With both the categorical and continuous approach, it is possible to compare children in particular groups (sociodemographic, geographic, etc.) in terms of whether they are over- or under-represented in the 'positive' or 'negative' outcome categories, as well as examining whether their average outcome scores differ. A sample of simple analyses is presented below to illustrate the diversity of ways in which the Index can be used. Because they are simple univariate or bivariate comparisons, and the data are cross-sectional, caution needs to be used in interpreting these findings. The analyses are given purely to be illustrative. The analyses here are based on weighted data.

Illustration of positive and negative indices (categorical form): Do metropolitan and ex-metropolitan infants differ in outcomes?

Figure 9 shows the number of children living in metropolitan areas ('met') and nonmetropolitan areas ('xmet') who are in the top 15% on 0, 1 or 2 domains (positive index, two left columns) and in the bottom 15% on 0, 1, 2 or 3 domains (negative index, two right columns). It indicates that there is a trend for more non-metropolitan infants to have strengths, and fewer to have areas of difficulty.

Figure 9: Scores on the categorical Outcome Index by metropolitan vs nonmetropolitan residence for infants.

Illustration of continuous form of Index: Does 4-year-old outcome differ by family income?

Figure 10 shows the average scores on the continuous Outcome Index scale for 4year-olds with family incomes ranging from less than \$499 per week to \$2000 or more. It reveals a modest linear gradient by income, from an average score of 96 for the lowest income group to around 104 for the highest group.

Figure 10: Scores of four-year olds on the continuous Outcome Index by family income

Illustration of continuous domain score: Is parenting self-efficacy related to 4year-olds' social and emotional development?

The primary parents of the 4-year-olds were asked to rate their own ability as a parent, on a 5-point scale as shown in Figure 11. About two-thirds of parents rated themselves as very good or better than average parents, and their children were on average above 100 on the continuous social-emotional scale. Children of parents who saw themselves as average had an average score of about 98. The small groups of parents who saw themselves as having some trouble or not very good at being a parent had children with markedly lower social-emotional scores.

Figure 11: Scores of children on the Social/Emotional domain continuous score by Parent 1's rating of their own parenting ability

Illustration of categorical domain score: Are there gender differences in learning outcome for 4-year-olds?

Figure 12 illustrates the use of the categorical form of a domain score, and shows that substantially fewer boys (11%) than girls (19%) were in the top 15% of the distribution on the Learning domain (positive index) and substantially more were in the bottom 15% (negative index) (20% boys versus 10% girls).

Figure 12: Percentages of boys and girls identified by the positive and negative learning domain cutoffs for four-year-olds.

Illustration of across-domain comparisons: Are infants with poor physical outcomes likely to have poorer social-emotional and learning outcomes?

This analysis examines the patterning of negative outcomes, using one categorical domain index (physical) and two continuous domain scores (for learning and social-emotional). Figure 13 shows that infants with poor physical outcomes had lower scores on both the learning and social-emotional domains than those who did not have a poor physical outcome.

Figure 13: Scores of infants on the Learning and Social/Emotional domains by whether they were identified by the negative Physical cutoff.

Conclusions

While it was initially intended to model the Outcome Index on the categorical approach adopted for the Canadian Vulnerability Index, further investigation on the implementation of this approach identified significant difficulties with it. Its inherent weaknesses included reliance on multiple arbitrary cut-offs and the impossibility of weighting different elements roughly equally. Most of these limitations were avoided with the continuous approach which is more statistically and mathematically principled. This approach was therefore adopted.

A strength of the Outcome Index as it is conceptualised here is that it provides not only the overall Outcome Index but also domain scores. It is thus possible to examine children who have problems on some domains but not all, and those with strengths in one domain and weaknesses in another, and so on. Thus, while the overall Outcome Index gives an overall summary, the domain scores account for the fact that development may not be uniform across domains. Further, the continuous and categorical forms of the Index allow flexibility in analytic uses of the Index.

The illustrative analyses suggest that the Index is working as would be predicted, and is capable of being discriminating between children with differing outcomes. The versatility of the Index is demonstrated in these examples.

Some limitations in the Outcome Index derived from Wave 1 data should be noted:

• *Discriminability is stronger at the problem end than the positive end:* Many variables in LSAC are designed to identify problematic or below-average child functioning, and their capacity to identify those with particular strengths or above-average functioning is often weak. Hence the final distribution has greater

negative discrimination than positive. Further, it was not possible to derive a meaningful index of positive physical outcomes for infants. Care is needed in the interpretation of the positive index.

- *Gaps in infancy data*: There are limited areas where it was possible or meaningful to collect 'outcome' information on the infant cohort.
- *Cut-offs are arbitrary*: The categorical form of the Outcome Index uses cut-offs to identify the top and bottom 15% of the distribution. There is no claim that these proportions are clinically meaningful. They are statistically based, in accord with the common view that one standard deviation below the mean of a population represents significant difficulty. It is therefore not possible to make general claims about the sample overall, such as "X% of children have low social competence" or "Y% of children are in excellent physical health" since the proportions in all cases are pre-defined. However, statements about subgroups of the sample relative to each other are possible. This is the prime purpose of the Outcome Index.

References

Australian Council for Educational Research (1999) "Who Am I?" Camberwell, Vic: ACER

Goodman, R. (1999) Strengths and Difficulties Questionnaire. See www.sdqinfo.com

- Sanson, A., Prior, M., Oberklaid, F., Garino, E. and Sewell, J. (1987) The structure of infant temperament: Factor analysis of the RITQ. *Infant Behaviour and Development*, *10*, 97-104.
- Prior, M., Sanson, A., Smart, D. and Oberklaid, F. (2000) *Pathways from infancy to adolescence: Australian Temperament Project 1983-2000.* Australian Institute of Family Studies, Melbourne.
- Varni, J. W. (1999) Measurement model for the Pediatric Quality of Life Inventory. *Medical Care*, 37, 126-139.
- Wetherby, A.M. & Prizant, .M. (2001) *Communication and Symbolic Behavior Scales – Infant/Toddler Checklist*. Paul H. Brookes Publishing Co.
- Willms, J. D. (Ed) (2002), Vulnerable children. Alberta: University of Alberta Press

Appendix A: Handling missing data

This appendix describes the procedures adopted to ensure that the presence of missing data did not skew results.

(a) 4-year-olds

In cases where one or more z-scores in a sub-domain were missing, a sub-domain score was still obtained by taking the average of all the available z-scores. However, when averaging, the standard deviation of the mean score increases as the number of scores averaged decreases. Hence children with more missing data for a sub-domain would tend to have scores further from the average value, without this being a reflection of their actual outcomes. To correct for this, a variable was calculated for each sub-domain with more than one variable, indicating the number of variables missing for each case. The frequencies of these variables can be seen in Table A1. These variables were used as grouping variables to divide the file by level of missingness for each sub-domain. A standard deviation score was then obtained for each level of missingness, which was used to divide the sub-domain score. This method of standardisation corrects for the greater standard deviation obtained when averaging fewer z-scores, without disguising any mean differences present in the data.

Table A1.	Numbers o	f variables	s missing for	r each	n sub-do	main	with	more t	than
one variab	le for four-	year-olds (unweighted	freq	uencies))			

	0	1	2
Health	4889	90	3
Social Competence	4969	0	
Externalising	4969	0	
Literacy	3177	63	1736
Approach to learning	3199	1728	

Note: For ease of interpretation frequencies indicating that all component variables of the sub-domain are missing are excluded from the above table.

It can be seen that there is no partially missing data for Social Competence and Externalising, but large amounts for the two variables relying on teacher report, and a moderate amount for Health. Due to insufficient numbers to reliably standardise scores, those with 1 and 2 missing variables on Health were merged, as were those with 1 and 2 missing variables in Literacy. Means and standard deviation scores for each sub-domain and level of missingness can be seen in Table A2. Note that post-standardisation the standard deviation for each sub-domain and each missingness level will be 1, while the mean will equal the prestandardisation mean divided by the prestandardisation standard deviation, e.g. poststandardisation the standard deviation will be 1. At this stage each sub-domain score was restandardised to have a mean of 0 and a standard deviation of 1 (ie a z-score) to correct for minor fluctuations from these values from taking averages and the standardisation process.

	0		1		
	Mean	SD	Mean	SD	
Health	.01	.63	44	1.11	
Literacy	.01	.77	05	.98	
Approach to learning	.03	.81	11	1.06	

Table A2. Means and standard deviations for pre-standardisation sub-domain scores for sub-domains which underwent standardisation by missingness level

The above approach repeated to standardise by missingness level when combining sub-domain scores into domain scores. The frequencies for the missingness variables for each domain can be seen in Table A3.

Table A3. Numbers of sub-domain scores missing for each domain for fouryear-olds (unweighted frequencies)

	0	1	2	3	
Physical	4198	784			
Social/Emotional	4968	1	0		
Learning	2895	1773	266	48	

Note: For ease of interpretation frequencies indicating that all sub-domain scores of the domain are missing are excluded from the above table.

Only one case was missing a sub-domain score for the Social/Emotional domain where other sub-domain scores were present; however a number of cases with some relevant data were missing sub-domain scores in the Physical and Learning Domains. Due to insufficient numbers to reliably standardise scores, those with 1 and 2 missing variables on Health were merged, and those with 3 missing variables in the Learning domain were grouped with those with 2. The means and standard deviations for the different levels of missingness for the Physical and Social/Emotional domains can be seen in Table A4. Scores at each missingness level were divided by their standard deviation.

Table A4. Means and standard deviations for pre-standardisation domain scores for domains which underwent standardisation by missingness level

	0	0		1		2	
	Mean	SD	Mean	SD	Mean	SD	
Physical	.01	.77	10	1.05			
Learning	.06	.72	09	.76	26	0.87	

Infants

There was only one measure for the Learning domain for infants so no adjustment for missingness was relevant. Table A5 shows the levels of missingness for the Physical and Social/Emotional domains. Given the low number of cases with 1 and 2 missing Social/Emotional subdomain scores, standardisation by missingness was not possible or necessary; therefore this was only done for the Physical domain.

Table A5. Numbers of variables missing for each domain with more than one variable for infants (unweighted frequencies)

	0	1	2
Physical	5029	77	
Social/Emotional	4311	1	2

Note: For ease of interpretation frequencies indicating that all component variables of the domain are missing are excluded from the above table.

As can be seen only the Physical domain had sufficient levels of partial missingness to require standardisation. The means and standard deviation of the prestandardisation Physical domain score for the two missingness levels can be seen in Table A6. Scores at each missingness level were divided by their standard deviation.

Table A4. Means and standard deviations for pre-standardisation domain scores for the physical domian

	0		1	
	Mean	SD	Mean	SD
Physical	.01	.77	85	1.22

Appendix B: Summary of variables included in the Outcome Index

0 year olds

Physical

Overall health rating: single parent-rated item of infants' health, from 1 (excellent) to 5 (poor).

Special health care needs: single derived yes or no item based on 6 component items indicating whether child needed medication or more health care than the average child due to a condition that has lasted or was expected to last 12 months or more

Social/Emotional

STSI Approach: mean of a 4-item parent-rated subscale from abbreviated Short Temperament Scale for Infants (STSI, Sanson et al., 1987; Prior, et al., 2000) on 1-6 frequency scale; low scores indicate withdrawing/shy, and high scores indicate approaching/sociable; response rate 84%; adequate reliability (alpha = .72) **STSI Irritability**: mean of a 4-item parent-rated subscale from abbreviated STSI, on 1-6 frequency scale; low scores indicate calm/not irritable, and high scores indicate irritable/volatile; response rate 84%; fair reliability (alpha = .57) **STSI Cooperativeness**: mean of a 4-item parent-rated scale from abbreviated STSI,

on 1-6 frequency scale; low scores indicate uncooperative, unadaptable, and high scores indicate adaptable, cooperative; response rate 84%; adequate reliability (alpha = .65).

Learning

CSBS Total: normed standardised score based on 24 items in the parent-report Communication and Symbolic Behaviour Scale (Wetherby & Prizant, 2001), items reflect the child's exhibition of various behaviours demonstrating emerging communication skills; good reliability (alpha = .89).

4 year olds

Physical

Health

Overall health rating: single parent-rated item of child's health, from 1 (excellent) to 5 (poor)

Special health care needs: single derived yes or no item based on 6 component items indicating whether child needed medication or more health care than the average child due to a condition that has lasted or was expected to last 12 months or more

Body-Mass Index (BMI): calculated from directly assessed variables of child's height and weight

Motor

PEDS QL Physical health subscale summary: 8-item parent-report Physical subscale from PEDS QL (Varni, 1999), largely assessing motor coordination, but also

containing 2 items about more general health, scaled to range from 0 (poor) to 100 (good).

Social/Emotional

Social competence

SDQ Prosocial: mean of 5 parent-rated items in the Prosocial subscale of the Strengths and Difficulties Questionnaire (SDQ, Goodman, 1999), assessing the child's propensity to behave in a way that is considerate and helpful to others, with items scored from 1 (Not true) to 3 (Certainly true); adequate reliability (alpha = .66). **SDQ Peer problems**: mean of 5 parent-rated items in the Peer subscale of the SDQ, assessing problems in the child's ability to form positive relationships with other children, with items scored from 1 (Not true) to 3 (Certainly true); fair reliability (alpha = .50).

Internalising

SDQ Emotional symptoms: mean of 5 parent-rated items in the Emotional Symptoms subscale of the SDQ, assessing a child's frequency of display of negative emotional states (e.g. nervousness, worry) with items scored from 1 (Not true) to 3 (Certainly true); fair reliability (alpha = .58).

Externalising

SDQ Hyperactivity: mean of 5 parent-rated items in the Hyperactivity subscale of the SDQ, assessing child's fidgetiness, concentration span and impulsivenesss with items scored from 1 (Not true) to 3 (Certainly true); good reliability (alpha = .74). **SDQ Conduct**: mean of 5 parent-rated items in the Conduct subscale of the SDQ, assessing child's tendency to display problem behaviours when interacting with others, with items scored from 1 (Not true) to 3 (Certainly true); good reliability (alpha = .69).

Learning

Language

PPVT: A standardised Rasch-modelled score based on interviewer administration of an abbreviated form of the Peabody Picture Vocabulary Test (PPVT-III Form IIA, 1997), a measure of receptive language;

Literacy

Parent rating of reading skills: parent rating on 3 yes/no items assessing whether a child has obtained reading skills at different levels of complexity, summed to give scores from 0 (good) to 3 (poor)

Teacher rating of reading skills: teacher yes/no ratings on 5 items; 0=poor skills, 5= strong skills; available for 64% of sample. Items assess the level of complexity a child is capable of reading as well as the child's interest in reading.

Teacher rating of writing skills: teacher yes/no ratings on 6 items; 0=poor skills, 6= strong skills; available for 65% of sample. Items assess the level of complexity of the child's writing skills as well as the child's interest in writing.

Numeracy

Teacher rating of numeracy skills: teacher yes/no ratings on 5 items; 0=poor skills, 5= strong skills; available on about 64% of sample. Items assess the child's ability to perform numeric tasks such as counting, classifying, and simple addition, along with the ability to recognise numbers.

Approach to learning

Who Am I? (WAI): Standardised score based on interviewer administration of the 'Who Am I?'(ACER, 1999), an Australian measure which assesses a child's ability to perform a range of tasks such as reading, writing, copying, and symbol recognition, as a measure of school readiness.