Development of an Index of Socio-Emotional Competence for Preschool Children in the Growing Up in New Zealand Study

Sahrish Ahmad[1]*, Elizabeth R. Peterson[1,2], Karen E. Waldie[1,2] and Susan M. B. Morton[2]

[1] Faculty of Science, School of Psychology, The University of Auckland, Auckland, New Zealand, [2] Growing Up in New Zealand, Centre for Longitudinal Research–He Ara ki Mua, The University of Auckland, Auckland, New Zealand

Growing international recognition of the importance of socio-emotional competence has led to calls to understand the prevalence and development of such competence in young children. However, socio-emotional competence is a multifaceted concept and hence difficult to track. The present research describes the processes we used to develop an index of socio-emotional competence in preschool children taking part in the Growing Up in New Zealand (GUiNZ) study. We draw on data collected from 6,156 children when they were 9 months, 2 years and 4.5 years of age. Twenty-six variables from seven measures were found to provide conceptual coverage of socio-emotional competence and had adequate discriminatory power and internal consistency. Our final sample consisted of 4,839 participants who had adequate data on the variables of interest. Exploratory factor analysis led to the emergence of three dimensions: easy-going, regulation, and exuberance. Multiple methods of weighting (equal, empirical, and time-based) were used to calculate the index, which resulted in eight possible index versions. The index with equally weighted variables (including observations scores) and dimension scores, and time weighted scores of socio-emotional competence at each data collection wave, was found to be the most empirically sound. Both categorical and continuous index scores were calculated to allow for variety of later analytical usage. Our findings may be of interest to policy-makers and clinicians who could use this knowledge to better understand the trajectories of development of socio-emotional competence, and factors associated with any derived patterns and change. This may also help identify children at risk of poor socio-emotional development as early as possible, which may be of interest to those looking to mitigate the risk of poor socio-emotional development.

Keywords: longitudinal, socio-emotional, composite index, parental report, weighting, preschool

INTRODUCTION

The aim of this paper is to describe the development of an index of socio-emotional competence, using data from a large national New Zealand prospective longitudinal study. The index draws on data collected during the first 5 years of child's life and includes key components of social-emotional development: emotional expression and understanding; regulation of emotions and behavior; social problem solving and social relationship skills (Rose-Krasnor, 1997; Denham, 2006).

Social-emotional competence is important not only because it helps children to learn (Denham, 2006; Britto, 2012) but also because it helps individuals establish and maintain healthy and meaningful relationships (Cohen et al., 2005). As the concept of social-emotional competence is multifaceted, a single index that draws on different components of socio-emotional competence provides an opportunity to try and summarize complex data for the ease of public, media, policy-makers, and researchers (Booysen, 2002; Joint Research Centre-European Commission, 2008). Our index seeks not only to identify an overall level of preschool socio-emotional competence, but also allows for the identification of groups of children with Low, Average and High socio-emotional competence at three data collection waves (DCWs: 9 months, 2 years and 4.5 years). This offers an opportunity to potentially track movement in and out of High and Low levels of socio-emotional competence across the preschool years.

Growing Up in New Zealand Study

The index was created using Growing Up in New Zealand (GUiNZ) data. The GUiNZ study began in 2008 and has been following the lives of ~7,000 children from the antenatal period (6 weeks before their birth) and plans to study them until their 21st year. Written informed consent was taken from all participating parents, and parents also provided consent on behalf of their participating child. The sample is representative of the New Zealand population in terms of key socio-demographic and ethnic characteristics. Recruitment was designed to ensure ethnic and socioeconomic diversity of the sample. At the 4.5-year assessment, the retention rate was 90 percent of the baseline (Morton et al., 2017). Full details about the design of the study can be found in Morton et al. (2014, 2017). The current paper details the conceptual framework we used to guide the development of our socio-emotional competence measure, methodological considerations in constructing an index, and how we calculated the index. Finally, we briefly explore how the index relates to pragmatic language and school readiness.

Conceptual Framework of our Socio-Emotional Competence Index

Socio-emotional competence involves an individual's ability to express, receive and manage emotions (Halberstadt et al., 2001; Denham, 2006; Clarke-Stewart and Parke, 2014) as well as their effectiveness in forming and maintaining relationships, and in general interactions (Rose-Krasnor, 1997; Han and Kemple, 2006). It also extends to include knowledge and abilities that an individual needs to make good life choices and deal with challenges (Leffert et al., 1997; Denham, 2006; Calaguas and Dizon, 2011).

As for the components of socio-emotional competence, Denham (2006) describes its core aspects as emotional expressiveness, understanding of emotions, regulation of emotion, and behavior, social problem solving and social relationship skills. These components correspond to the lowest level of a prism model of social competence described by (Rose-Krasnor, 1997), and each is reported to develop and operate interactively (Denham, 1998; Denham et al., 2012).

In creating our index, we aimed to tap into each one of these five aspects of socio-emotional competence to provide age-appropriate coverage. However, it is important to note that socio-emotional competence does not emerge all at once; developing as the child becomes more able to express themselves and interact socially. For example, new-borns are sensitive to correct facial stimuli (Beauchamp and Anderson, 2010) and are able to express their basic emotional states (Halberstadt et al., 2001). During the first year of their lives, infants also learn to use emotions instrumentally to influence their social environments (Denham, 2006). In toddlerhood, children start to develop the ability to regulate their emotions (Kopp, 1989; Shatz, 1994), while late preschool years witness further improvement in emotional expression (Denham et al., 1992); emergence of ability to understand the complexity of emotions (Harris, 1989; Kestenbaum and Gelman, 1995; Denham, 1998); and to assign meanings to emotions (Baron-Cohen et al., 1999; Denham, 2006).

Composite Indices of Children's Development

Composite indicators of child development combine data on singular indicators to give comprehensive information on children's functioning. While there are numerous measures available that measure a particular aspect of children's development, they do not help us gain a holistic picture of it which is needed to: calculate prevalence in different populations, and evaluate the impact of any policies designed to promote development and monitor progress toward any desired societal goals (Land et al., 2007). Composite indices also potentially allow for easier interpretation of the data and facilitate communication with the public as they allow users to compare diverse information in an effective manner (Joint Research Centre-European Commission, 2008).

While composite indexes have numerous advantages, they are not without their problems. Selection of the indicators and weights can be disputed. In addition, the composite index may oversimplify a complex phenomenon resulting in information being lost or wasted. For example, it may disguise deficiency in a domain; giving misleading information, and leading the policy makers to draw incorrect conclusions. In addition, inappropriate policies may also result if difficult to measure domains are ignored (Joint Research Centre-European Commission, 2008). As a result, care is always needed when interpreting a composite index.

Approaches to Measurement

There are numerous ways of developing a composite index. Typically, creating a composite index involves four steps: variable selection, scaling, weighting and aggregation (Booysen, 2002). These steps are not necessarily undertaken in sequential order, as the researcher might go back to change or re-scale the selected variables or readjust the assigned weights. Any decisions made when constructing an index should be made based on both the available data and the literature (Booysen, 2002; Joint Research Centre-European Commission, 2008; Cowan et al., 2012).

Variable Selection

Booysen (2002) in his review of 20 developmental indices argued that selection is a two-level process. The first step concerns how many and what domains make up the index. This selection of domains for the model should be guided by theory, evidence (gained through empirical analysis), and practicality and/or intuitive appeal (Diener and Suh, 1997). With respect to empirical evidence, variables more strongly associated with each other should be picked from the pool of applicable variables as the components are expected to be interdependent as they represent the same underlying developmental phenomena (McGranahan et al., 1972; Field, 2013). Field (2013) also states that the magnitude of the correlation does not need to be taken into account.

With regard to variable selection in psychology-related developmental indices, different criteria have been employed. For example, Sanson et al. (2005) and Misson et al. (2011) developed an index of well-being (including the domain of social and emotional understanding) using data from the Longitudinal Study of Australian Children (LSAC). They included variables providing good conceptual coverage without redundancy; with high response rates; ability to discriminate among good, average and poor performers (with 5–20% of the sample identified as good or poor); and adequate internal consistency for the variables included in the index. On the other hand, Williams et al. (2014), using data from the Growing Up in Ireland study, developed an index of deprivation, which also included emotional well-being. They prioritized keeping one variable in each domain while also trying to ensure that the selected variables provided sufficient conceptual coverage for their multi-dimensional deprivation index for 9-year-olds.

Scaling

After variable selection, the variables need to be scaled. According to Booysen (2002), already scaled data can be left untouched. Some studies use standardized variables while others use ordinal or conventional linear scaling. Booysen raises the concern that standardizing a distribution with outliers may give biased results, however, with large samples, the impact of (much less frequent) outliers is minimal. Several large-scale studies have scaled individual scores to z-scores to combine variables across the domains. For example, Blakemore and Gibbings (2006) developed a measure of socio-economic position of family by merging standardized data from the LSAC study and the Income and Labor Dynamics in Australia Survey (HILDA); and Sanson et al. (2005) and Misson et al. (2011) transformed variables for developing the LSAC outcome index of well-being.

Weighting and Aggregation

Weighting should be reflective of the relative importance of each variable (Drewnowski, 1974). There is support for equal weighting with burden of proof falling on the use of differential weighting (Booysen, 2002). For example, for large-scale studies like Sanson et al. (2005), Misson et al. (2011), and Blakemore and Gibbings (2006), equal weighting was carried out. This seems to be the most common approach with 14 out of 20 development indices reviewed by Booysen also employing equal weighting.

Differential weighting can be carried out by consulting with experts, which has been the conventional practice, or guided by principal component analysis/factor analysis, in which variables and/or the components/factors are weighted in accordance with the variance they explain (Hollingshead, 1975; Ram, 1982; Slottje, 1991; Nicoletti et al., 2000; Booysen, 2002; Joint Research Centre-European Commission, 2008). While weighting based on experts' opinions can be biased, use of empirical methods can also be too stringent although methodologically defensible. Wish (1986) calls for a balance between subjective and objective weighting.

After the assignment of weights, aggregation begins, which can be functional or additive (Booysen, 2002). The former entails combining the variables together based on their functional relationships with each other. Composite indices also require validation, which can be internal (item-analysis) or external (regression against a theory-supported outcome variable). Validation can also lead to the need to make subsequent adjustments in selection, scaling, weighting and aggregation (Ul Haq, 1995; Booysen, 2002).

CALCULATION OF THE SOCIO-EMOTIONAL COMPETENCE INDEX

Statistical Analysis

The analyses for the present study were performed using IBM Statistical Package for Social Sciences (SPSS; Version 22) and IBM SPSS AMOS (version 25). Exploratory Factor Analysis (EFA) and bivariate analyses were conducted on SPSS while Confirmatory Factor Analysis were carried out on AMOS.

Selection of Variables

This study included data collected on children when they were 9 months ($N = 6,476$), 2 years ($N = 6,327$), and 4.5 years ($N = 6,156$) of age. In keeping with literature cited, we initially identified all the variables across the data set that provided good conceptual coverage (age-appropriate) to socio-emotional competence. This index was theorized to include domains of: emotional expression, understanding, regulation of emotions and behavior, social problem solving and social relationship skills (Rose-Krasnor, 1997; Denham, 2006). In addition, internal consistency of the tools and discriminatory power (ability to differentiate between poor and good performers) of the variables was considered, with poor and good performers each accounting for 6–29 percent of the sample.

One criterion for variable selection is that they must be interdependent (Field, 2013). For the present study, the variables that we included in the index were all significantly associated with all, or at least most, of the other variables. In our study, the variable that correlated with the least number of other variables, was still associated with 65% of the other variables. We could not fully meet the criterion of only including variables that were correlated with all the other variables as it would have caused several variables to be dropped. In addition, as our variables were measured at three different DCWs spanning a 4 year period of rapid child development, it is arguably unrealistic to expect the variables measured early to correlate with all the other variables

in the pool across time. Hence, in this study, we kept the variables that correlated with most, if not all, the other variables.

Many of the socio-emotional competence related measures selected comprised of subscales which have multiple items. As the components of socio-emotional development are expected to operate interactively, many of the variables were expected to measure more than one aspect of socio-emotional competence. **Figure 1** shows the expected categorization of the selected variables based on how they are usually treated in the literature. In some case, the variables could be argued to fall under more than one component of socio-emotional competence; these are italicized and in parentheses.

The final index of socio-emotional competence was derived from the following measures administered to the GUiNZ cohort. These are listed below, in time order (see also **Figure 1**).

Infant Behavior Questionnaire–Revised Very Short Form (IBQ-R VSF)

The IBQ-R VSF (Putnam et al., 2014) measures temperament in infants between the ages of 3–12 months. Temperament is argued to be a relatively stable and largely biological aspect of an individual's early personality, which involves individual differences in reactivity and self-regulation. Reactivity refers to the "arousability" of emotional, attentional and motor responses, while self-regulation involves modulation of reactivity (Rothbart, 1981). These two components of temperament relate to emotional expression and regulation of emotions and behaviors, which are core parts of our socio-emotional competence model The IBQ-R VSF is a 36-item scale which asks parents to comment on the frequency of certain behaviors, with responses ranging from "1 = *Never*" to "7 = *Always*" on a seven-point Likert scale. The additional responses of "*Does not apply,*" "*Don't know,*" and "*Refused*" included in GUiNZ were coded as missing. The responses are later averaged to calculate scores across each temperament factor.

The original IBQ- R VSF was designed to identify three temperament factors: Negative Emotionality, Positive Affect/Surgency and Orienting/Regulatory Capacity (Putnam et al., 2014). However, Peterson et al. (2017a), using the current study's cohort, found that the three-factor structure of the IBQ-R-VSF had poor model fit when administered to 5,639 mothers of infants aged 23 to 54 weeks. Instead, they found evidence for a five-factor structure including broadly the original three factors: Negative Emotionality, Positive Affect/Surgency and Orienting Capacity (revised from Orienting/Regulatory Capacity), and two new factors: Affiliation/Regulation and Fear. The five-factor model demonstrated acceptable model fit on two randomly created samples of more than 2,300 participants: sample 1, $\chi^2_{(619)}$ = 6,384; χ^2/df = 10.31, $p < 0.001$; RMSEA = 0.06; CFI = 0.77; gamma hat = 0.99; SRMR = 0.06; Sample 2, $\chi^2_{(619)}$ = 6,355; χ^2/df = 10.27, $p < 0.001$; RMSEA = 0.06; CFI = 0.77; gamma hat = 0.99; SRMR = 0.06 (Peterson et al., 2017a). While the CFI was low for both samples, this is not surprising as the CFI is sensitive to complex models (Fan and Sivo, 2007). In addition, since both gamma hat and SRMR (model fit indices considered to be the most stable (Fan and Sivo, 2007) were both good, the overall
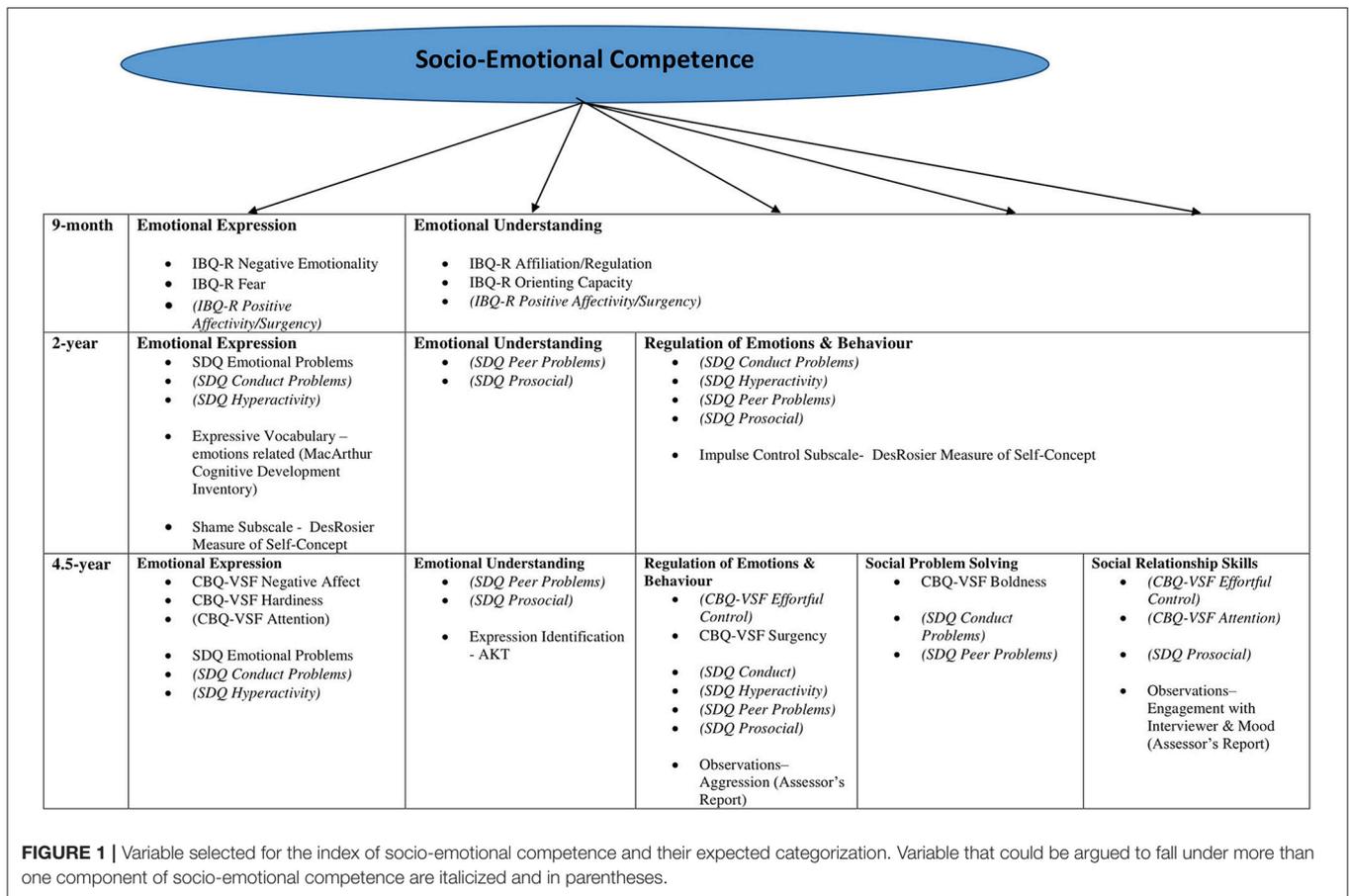
model fit was deemed good. In accordance with Tabachnick and Fidell (2001), the Cronbach's alpha reliabilities of the five temperament factors were also found to be acceptable (range = 0.63–0.89). The measure was also reported to be similarly precise across the four major ethnic groups included in GUiNZ's cohort (Peterson et al., 2017b). In the current study, the Peterson et al. revised five-factor structure was used.

Strengths and Difficulties Questionnaire (SDQ)

SDQ (Goodman, 1997) is widely used for assessing mental health problems and psychological adjustment in children and adolescents. It consists of five scales: emotional problems, behavioral problems, hyperactivity, peer problems, and prosocial behavior (Goodman, 1997), which relate to all aspects of socio-emotional competence as outlined by Denham (2006). It is a 25-item long measure, which can be completed by parents and teachers of children aged 2–16 years with responses including: "0 = *Not True*," "1 = *Somewhat True*" and "2 = *Certainly True.*" GUiNZ included other response options of "Don't Know" and *Refused*," which were coded as missing. Scores for each factor are summed. (Goodman, 1997; Youth in Mind, 2014). GUINZ used pre-school SDQ (Youth in Mind, 2014) at 2-year and the standard version at 4.5-year DCWs.

The SDQ has been successfully used across diverse populations with varying socioeconomic levels (Woerner et al., 2004; Achenbach et al., 2008). Furthermore, satisfactory reliability with Cronbach's Alpha range of 0.71–0.84 (according to the standard given by Nunnally et al., 1967) for all scales was reported. The exception to this was the peer problems scale which had a lower Cronbach's alpha reliability α =0.54 at 2 years in the GUiNZ cohort. This may reflect the social developmental stage of the children, and also the difference in the interpretation of its underlying construct by the raters for 2-year olds. Note also, the preschool version of SDQ has not been used much in comparison to the standard SDQ version (D'Souza et al., 2017).

The factor structure of the SDQ has also been a subject of discussion. The original five factor structure has been supported for a representative sample of adolescents (13–18 years old) in the US (He et al., 2013) and young children (3, 5, and 7 years old) in Britain (Croft et al., 2015). There has also been support for using a three-factor structure with internalizing subscale (comprising emotional and peer problems scales), externalizing subscale (comprising behavioral problems and hyperactivity scales) and a prosocial subscale, for a representative low risk sample of children (5–16 years old) in Britain (Goodman et al., 2010) and for a large sample of children (5–7 years old and 10–12 years old) in Denmark (Niclasen et al., 2013). For GUiNZ, D'Souza et al. (2017) found support for a modified five-factor model in 2 year olds, in which the prosocial factor extended into a positive construal factor by including cross-loadings of reverse-coded items. However, D' Souza et al. suggested the need for their findings to be replicated with other children within the age group, and for the SDQ scoring structure to be modified if results consistent with their findings are found. In the meantime, in the absence of procedures for computing an index or composite scores with cross-loading variables, we used the original five factor structure (Goodman, 1997) in the present study.

FIGURE 1 | Variable selected for the index of socio-emotional competence and their expected categorization. Variable that could be argued to fall under more than one component of socio-emotional competence are italicized and in parentheses.

DesRosier's Measure of Self-Concept (1990)

DesRosier's measure of self-concept is a multidimensional tool that gives scores on perceptual self-recognition, cognitive self-representation, linguistic self-description, motivational self-assertion, emotional self-evaluation, and social self-regulation (DesRosiers and Busch-Rossnagel, 1997). It has a four-point response set ranging from "1 = *Not at all typical of my child*" to "4 = *Very typical of my child,*" with additional options of "*Don't know,*" and "*Refused*" coded as missing.

The self-regulation subscale was included in the present study as it relates to the regulation of behavior and emotions aspects of socio-emotional competence. However, inspection of the seven items of the scale showed that they grouped conceptually under two themes. Four items are linked to the experience of guilt and shame before and after doing something wrong: "Is upset, ashamed or sorry when [he/she] shows you [he/she] has done something bad;" "Is upset, ashamed or sorry when you find [him/her] doing something [he/she] should not;" "Calls your attention to things [he/she] did that [he/she] was not supposed to do;" and "Tries to hide something [he/she] was not supposed to do," are examples of this. The remaining three items relate to controlling impulses: "Stops [himself/herself] from doing something [he/she] wanted to do because you were watching;" "Stops [himself/herself] from doing things you have said may hurt;" and "Waits for things [he/she] has been told to wait

for." Therefore, this subscale of socio-emotional competence was partitioned into two groups with Confirmatory Factor Analysis (performed using IBM SPSS AMOS Version 25) showing the following acceptable model fit: $\chi^2 = 401.443$; $\chi^2/df = 30.88$, $p < 0.001$; RMSEA $= 0.079$; CFI $= 0.93$; gamma hat $= 0.97$; SRMR $= 0.05$. While, a borderline value was observed for RMSEA (Fan and Sivo, 2007), Fan and Sivo, advocate for greater emphasis being placed on gamma hat and SRMR values as they are not sensitive to model complexity and model misspecifications. In our model these values provided support for excellent model fit (Hooper et al., 2008).

Emotion Words From the MacArthur Cognitive Development Inventory

In order to assess expressive vocabulary, the MacArthur Cognitive Development Inventory - Toddler Short Form (Fenson et al., 2000) was used at the 2-year DCW with a yes or no response. The tool asks parents to report on whether a child uses particular words that are age appropriate, such as dog, mother etc., to gauge language development in infants and toddlers. The shorter form used in GUiNZ consists of 101 items evaluating their use in six languages: English, Maori, Samoan, Tongan, Cantonese, and Mandarin. For the current study, words such as uh-oh, ouch, friend, like, hug, gentle, wish, and happy were extracted to provide an indication of emotion knowledge both

internal and external. Since these words are adopted for Samoan, Tongan and Mandarin, and directly translated for Maori; we only used scores for English and Maori. Children were able to attain a maximum of 1 score for usage of each word (that is, they were not given a score of 2 if they knew the word in both the languages). The scores were then summed. This is an internationally recognized, reliable, valid, cost and time efficient tool (Fenson et al., 2000), however, our extraction of items related to emotions is novel.

Child Behavior Questionnaire–Very Short Form (CBQ-VSF)

The CBQ-VSF (Putnam and Rothbart, 2006) is a measure of temperament in children aged between three 3 and 7 years. It was used to assess temperament of children when they were ~4.5-years old. The CBQ very short form consists of 36 items across three broad scales of surgency (defined by activity level, high intensity pleasure, and impulsivity), negative affect (defined by sadness, fear, discomfort, and anger/frustration) and effortful control (defined by low intensity pleasure, inhibitory control, and attentional control). The inventory asks the caregivers to report on the level of different child behaviors using a seven-point Likert scale range from "1 = *Extremely untrue*" to "7 = *Extremely t*rue." The GUiNZ also included additional response options of "*Not applicable*," "*Don't know*", and "*Refused*," which were coded as missing.

CBQ-VSF reliability does not vary by ethnicity, location, socioeconomic status and age, and shows adequate validity levels (Hughes et al., 2008; de la Osa et al., 2013). When applied to GUiNZ data, Stubbing et al. (2017) found that a three-factor structure was not the best fit of the data, and that a six-factor structure (adding hardiness, fear, and attention to the original three factors) was the most parsimonious and context-sensitive. Confirmatory Factor Analysis (CFA) found that the six actors had an acceptable fit, $\chi^2 = 11670.911$; $\chi^2/df = 21.42$, $p < 0.001$; RMSEA = 0.06; CFI = 0.71; gamma hat = 0.90; SRMR = 0.07. Again CFI value was arguably low; however, since CFI is sensitive to complex models, and gamma hat and SRMR (which are stable model fit indices) were within acceptable ranges for the model (Fan and Sivo, 2007), this model fit was accepted. The Cronbach alpha reliabilities of five of the six factors (Negative Affect, Fear, Effortful Control, Surgency and Attention) were found to be acceptable (range = 0.61–0.72). The Cronbach alpha reliability for Hardiness was just below the acceptable range with value of.58 (Tabachnick and Fidell, 2001). This six-factor structure was used for the current study.

Affective Knowledge Task (AKT)

The AKT (Denham, 1986) is a widely used measure tapping into knowledge of emotions (Morgan et al., 2009), which is essential to socio-emotional competence in pre-schoolers. In the GUiNZ study, only the expression identification task was used due to time constraints at 4.5-year DCW. A few other modifications were also made. First, rather than using puppets, cards with simple cartoon faces showing different emotions were used. In order to avoid potential ceiling effect, which starts to appear at around 4.5 years (Denham, 2006), two more emotions of surprise

and disgust were added to the original four emotions tested. In addition, eyebrows were removed from the faces to make them more gender-neutral. In the current study, a score of "2"was given when the child named the correct emotion or its appropriate synonym; "1" when the child named the incorrect emotion, but it had the same emotional valence as the correct response, and "0" for when the child gave the incorrect emotion and it was of the opposite emotional valence.

The AKT has been found to have good internal consistency and test-retest stability over a 1-year period (Denham, 2006; Denham et al., 2012). It is also argued to be ecologically valid owing to the small amount of verbalization required during the administration and the fact that it can be performed during play (Denham, 2006). For GUiNZ data, it shows good variability (Morton et al., 2017) signifying that the addition of emotions of surprise and disgust resolved the problem of a potential ceiling effect in the cohort.

Assessor Report From Preschool Self-Regulation Assessment (PSRA)

The PSRA (Smith-Donald et al., 2007) consists of tasks administered individually to children to assess their self-regulatory skills in emotional, attentional and behavioral domains (Roid and Miller, 1997; Wakschlag et al., 2005). The Assessor Report, which comprises 28 items, is part of this battery. The report consists of an observer describing the child's emotions and behaviors during the PRSA assessment. Aside from that, these items provide an indication of issues that may have influenced the child's performance during data collection, as well as an avenue for comparison between parental reports and observations. The GUiNZ used 13 items from the Assessor Report. For the present study, we focused on questions related to the child's attempt to engage with the interviewer ("Actively attempts to engage interviewer"); mood observed ("Child shows intense apprehensive, sad, or worried feelings during session"), and aggression (verbal, physical and toward objects). Items related to child engagement with the interviewer and mood had four response options and the three questions related to aggression were dichotomous. The latter three items were summed together to give a single score of aggression. Further scoring of these items is described below in the scaling section.

Missing Data

Out of the 6,156 participants, who had data across all three DCWs, some had incomplete data on the selected socio-emotional competence related measures. Therefore, with the exception of IBQ-R VSF, entries for participants ($n = 1317$, 21.4%) having more than 10 percent of the data missing on one or more selected measures were deleted. For the IBQ-R VSF, 10.8 percent was used as a cut off as a 10 percent cut-off would have resulted in deletion of 336 (5.5%) more cases.

Participants with more than 10 percent missing data (10.8% in the case of IBQ-R VSF) did not differ significantly from the remaining participants in terms of gender, parity and mother's age. However, there was a significant difference between both groups of participants in terms of mother's ethnicity and socio-economic status (with mothers with European ethnicity and

higher socio-economic status overrepresented in the imputed sample). The difference was, however, not large: Cramer's V < 0.07 (Cohen, 1988). See **Table 3**, for demographics of the full and restricted sample.

Expectation Maximization (EM) was used to impute the missing values that remained after deletion of cases with more than 10 to 10.8% percent missing values. This missing data procedure was used as it runs the model till no possibility of statistically significant improvement in model fit remains. It is based on the premise that the observed data is the best estimation of the missing data. It uses the available (observed) data to calculate a value (estimation) of the missing value using formulae that are meant to predict the variables the best (Dempster et al., 1977), and then checks if the value is most likely. If it is not, then values are re-imputed. EM is not biased like mean substitution or regression substitution. It does not affect the correlation values between variables, making it an optimal procedure for regression or factor analysis (Schafer and Olsen, 1998).

An important assumption of EM is that the data is missing completely at random or missing at random. Little's MCAR test was provided by the EM procedure for scales: IBQ-R- VSF (χ^2 = 34909.20, df = 30,219, p < 0.001); SDQ at 2 years (χ^2 = 2588.30, df = 2,025, p < 0.001); Desrosier self-concept measure (χ^2 = 6301.10, df = 5,050, p < 0.001); SDQ at 4.5 years (χ^2 = 1,210, df = 855, p < 0.001) and CBQ-VSF (χ^2 = 3310.19, df = 2,572, p < 0.001). However, this chi-square test is extremely sensitive to large sample sizes (Tanaka, 1987). Therefore, chi square ratios to df were calculated for each scale: IBQ-R VSF (χ^2/df = 1.16, p = 0.281); SDQ at 2 years (χ^2/df = 1.28, p = 0.258); DesRosier self-concept measure (χ^2/df = 1.25, p = 0.264); SDQ at 4.5 years (χ^2/df = 1.42, p = 0.233) and CBQ-VSF (χ^2/df = 1.29, p = 0.256). These were all non-significant, suggesting that the missing values are missing completely at random; thus, lending further support to the use of the EM method of imputation to estimate the missing data.

The imputed sample comprised of 4,839 participants, which did not differ significantly from the original representative sample with regards to child's gender, parity (at the antenatal DCW), mother's age and socio-economic status (at the antenatal DCW). Significantly more children with mothers of European ethnicity (self-prioritized) were included in the imputed sample than in the original sample at the antenatal period. However, the difference, though significant, was not large: Cramer's V < 0.04 (Cohen, 1988).

Scaling

To scale scores for the composite index, scores must be on an interval scale (Field, 2013). Scores for the vocabulary task (2-year DCW) and Affective Knowledge Task (4.5-year DCW) were already on an interval scale. The IBQ-R VSF, SDQ, and CBQ-VSF with Likert scales were also treated as interval scale. Scaling observations related items was not straightforward. There were two Likert-like items used to measure engagement and mood observed in the Assessor's Report at 4.5-years, with four possible response options and three dichotomous items related to aggression.

While Likert items have rank order, the difference between the values cannot be presumed to be equal (Sullivan and Artino, 2013). For example, the distance between the response options "Child frequently initiates conversation by asking questions, sharing information", "Child initiates conversation on occasion and is responsive to interviewer through eye-contact, talking, or smiling", and "Child does not initiate conversation, is slow to warm up," cannot be stated to be equal. Therefore, consideration of Likert-item or scale as an interval scale has been controversial (Knapp, 1990; Jamieson, 2004). Norman (2010) and Brown (2011) also highlighted the difference between a Likert-item, which is usually a single statement/question with a response-set consisting of four or more points and Likert-scale, which is an aggregate of responses on Likert-items. They noted that researchers have often used the terms Likert scale and Likert item interchangeably; however, their meanings are discrete, and they provided evidence that Likert scales can be considered continuous.

In the current study, the creation of a composite score for the five items from the Assessor's Report was considered. Exploratory factor analysis (EFA) with Direct Oblimin rotation using Principle Axis Factoring (PAF) showed that these items loaded well on two factors with loadings ranging from 0.52 to 0.53 for two items related to mood and engagement, and loadings ranging from 0.39 to 0.67 for three items related to aggression. These items grouped into two factors: observations related to mood and engagement of the child (α = 0.36), and aggression (α = 0.55). The low alpha values of these factors could be attributed to the short length of the scales. These variables were included in the analyses in spite of the low reliability as observations-based data adds breadth to the data (Carter et al., 2004). Finally, z-scores were calculated to scale all variables used in the analysis.

Weighting and Aggregation

Following the selection and scaling of the variables, EFA with Direct Oblimin (Oblique) rotation was performed with variables at each DCW to see how the variables grouped together. It was decided to keep variables that have a minimum factor loading of 0.162 (Stevens, 2002; Field, 2013). According to Stevens, with larger data sets (as it is the case with GUiNZ), adequate factor loading can be set at the minimum of 0.162. In addition, it was decided to use Principal Axis Factoring (PAF) rather than Maximum Likelihood (ML) and Principal Component Analysis (PCA), as the former is better for data which is not normally distributed (and our variables were not all normally distributed, see **Table 2**), and unlike PCA, it also takes into account individual variance. Oblique rotation (which assumes at least some correlations between factors) was chosen because in keeping with our selection criteria, the socio-emotional variables making up the index were required to be correlated with each other (Costello and Osborne, 2005). Costello and Osborne note that if the factors are uncorrelated, both Orthogonal and Oblique rotations give almost identical results.

Following EFA, three dimensions emerged: easy-going, regulation and exuberance. The aggression observations scale

TABLE 1 | Summary of exploratory factor analysis results for variables at the three data collection waves.

Variables	Factor loadings		
	Factor 1 (Easy-going-low)	Factor 2 (Regulation)	Factor 3 (Exuberance)
9 MONTHS			
Negative emotionality–IBQ-R VSF	**0.840**		
Fear–IBQ-R VSF	**0.443**		
Orienting capacity–IBQ-R VSF		**0.617**	
Affiliation/ Regulation–IBQ-R VSF		**0.612**	
Positive affectivity/ Surgency–IBQ-R VSF	0.220	**0.422**	
Eigenvalues	1.59	1.45	
% of variance	31.84	28.93	
2 YEARS			
Conduct problems–SDQ	**0.676**		
Emotional problems–SDQ	**0.548**	0.257	
Hyperactivity problems–SDQ	**0.521**		
Peer problems–SDQ	**0.511**		
Prosocial–SDQ	−0.361	**0.322**	
Expressive vocabulary–emotions related	**−0.277**		
Impulse control Subscale– DesRosier's Measure		**0.642**	
Shame subscale– DesRosier's Measure		**0.596**	
Eigenvalues	2.25	1.56	
% of variance	28.08	19.51	
4.5 YEARS			
Negative affect–CBQ VSF	**0.767**	1.65	
Emotional problems–SDQ	**0.620**		−0.357
Hyperactivity problems–SDQ	**0.581**	−0.314	−0.389
Conduct problems–SDQ	**0.577**	−0.238	
Peer problems–SDQ	**0.396**		−1.85
Hardiness–CBQ VSF	**0.264**		
Effortful control–CBQ VSF		**0.705**	
Attention–CBQ VSF		**0.523**	
Prosocial–SDQ		**0.492**	

(Continued)

TABLE 1 | Continued

Variables	Factor loadings		
	Factor 1 (Easy-going-low)	Factor 2 (Regulation)	Factor 3 (Exuberance)
Surgency–CBQ VSF			**0.616**
Fear–CBQ VSF			**0.546**
Observations– engagement with interviewer and mood (Assessor's Report)			**0.381**
Expression identification			**0.184**
Eigenvalues	2.78	1.52	1.95
% of variance	21.40	11.68	15.03

Factor Loadings <0.162 are not shown; factor 1 recovers a low Easy-going dimension, this is subsequently inverted in future analysis so that each dimension has a positive valence.
Items are bolded to show the factor under which each variable was placed for subsequent analysis.

did not load on any of the three dimensions generated in the EFA so it was dropped from the analysis. The third factor (exuberance) only emerged at the 4.5-year DCW (refer to **Table 2** for how the variables were grouped) as the childrens' behavior starts to become more developmentally complex (Rothbart, 2011; Campbell et al., 2016).

Of note, the three dimensions that emerged did not correspond directly with the five components of socio-emotional competence theorized by Denham (2006). This reflects the measurement tools we selected for the index. In selecting our measures, we aimed to get age appropriate coverage of the five aspects of socio-emotional competence described by Denham (2006). However, most of the measures we used tapped into more than one of the five components of socio-emotional competence and hence, while the index covered the model, it did not individually assess its components. For example, the peer problems scale of SDQ at the 4.5-year DCW assessed emotional understanding, emotion regulation and social relationship skills, while the conduct problems scale tapped into emotional expression, social problem solving and regulation. Similarly, attention from CBQ VSF assessed emotional expression and social relationships. Therefore, as our measures generally measured more than one component of socio-emotional competence, rather than the EFA identifying separate discrete components of socio-emotional competence, higher-order dimensions (comprising of more than one component given in the socio-emotional competence model) were recovered. It is also worth mentioning that the components of socio-emotional competence do not operate discretely in the real world. They are constantly interacting with each other, and what we see or assess in the real world is a dynamic combination of these components.

The three dimensions we identified (easy-going, regulation and exuberance) share some similarity with the child behavior profiles of "difficult," "easy/regulated" and "slow to warm up,"

TABLE 2 | Showing the grouping of socio-emotional variables after exploratory factor analysis at each DCW and related descriptive statistics.

Cohorts	Easy-going	Regulation	Exuberance
9-month	• IBQ-R VSF negative emotionality ($M = 3.38$, $SD = 1.06$, Skew = 0.33, Kurt = −0.43) • IBQ-R VSF fear ($M = 3.29$, $SD = 1.80$, Skew = 0.53, Kurt = 0.92)	• IBQ-R VSF orienting capacity ($M = 4.49$, $SD = 1.05$, Skew = 0.04, Kurt = −0.48) • IBQ-R VSF affiliation/regulation ($M = 5.70$, $SD = 0.77$, Skew = −0.74, Kurt = 0.85) • IBQ-R VSF positive affectivity/surgency regulation ($M = 5.17$, $SD = 0.71$, Skew = −0.32, Kurt = 0.22)	
2-year	• SDQ conduct problems ($M = 3.09$, $SD = 1.97$ Skew = 0.54, Kurt = −0.08) • SDQ emotional problems ($M = 1.77$, $SD = 1.58$, Skew = 1.24, Kurt = 1.83) • SDQ hyperactivity ($M = 4.31$, $SD = 2.11$, Skew = 0.20, Kurt = −0.31) • SDQ peer problems ($M = 2.12$, $SD = 1.63$, Skew = 0.65, Kurt = 0.00) • Expressive vocabulary–emotions related (MacArthur Cognitive Development Inventory) ($M = 4.18$, $SD = 2.30$, Skew = −0.09, Kurt = −1.00)	• Impulse control subscale- DesRosier's measure ($M = 3.38$, $SD = 1.06$, Skew = −0.20, Kurt = −0.15) • Shame subscale - DesRosier's measure of self-concept regulation ($M = 3.38$, $SD = 1.06$, Skew =0.00, Kurt = −0.40) • SDQ prosocial regulation[b] ($M = 3.38$, $SD = 1.06$, Skew = −0.41, Kurt = −0.17)	
4.5-year	• CBQ-VSF negative affect ($M = 4.37$, $SD = 0.89$, Skew = −0.09, Kurt = −0.06) • SDQ emotional ($M = 1.94$, $SD = 1.77$, Skew = 1.11, Kurt = 1.08) • SDQ hyperactivity[a] ($M = 3.89$, $SD = 2.27$, Skew = 0.38, Kurt = −0.34) • SDQ conduct problems ($M = 2.23$, $SD = 1.70$, Skew = 0.63, Kurt = 0.21) • SDQ peer problems ($M = 2.27$, $SD = 1.56$, Skew = 1.06, Kurt = 0.81) • CBQ-VSF hardiness ($M = 3.96$, $SD = 1.14$, Skew = −0.02, Kurt = −0.38)	• CBQ-VSF effortful control ($M = 5.44$, $SD = 0.66$, Skew = −0.48, Kurt = 0.78) • SDQ prosocial ($M = 5.73$, $SD = 1.01$, Skew = −0.61, Kurt = −0.17) • CBQ-VSF attention ($M = 7.83$, $SD = 1.75$, Skew = −1.16, Kurt = 1.76)	• CBQ-VSF fear ($M = 3.92$, $SD = 1.11$, Skew = 0.12, Kurt = −0.15) • CBQ-VSF surgency ($M = 5.21$, $SD = 0.81$, Skew = −0.56, Kurt = 0.42) • Observations– engagement with interviewer & mood (Assessor's Report) ($M = 6.94$, $SD = 7.95$, Skew = −1.17, Kurt = 1.99) • Expression identification - AKT ($M = 1.02$, $SD = 2.140$, Skew = −1.00, Kurt = 2.05)

[a],[b]Prosocial subscale (2-year DCW) and Hyperactivity subscale (4.5-year DCW) cross-loaded on two dimensions: easy-going and regulation child.

respectively, given by Thomas et al. (1970), who longitudinally assessed 141 children from infancy to pre-school and elementary school years using parental ratings complimented by observations. The "difficult" profile was characterized by a child who cries a lot, withdraws from new situations, has irregularities in bodily functions and throws tantrums. The "easy" profile included positive mood, adaptability, regularity in bodily functions and low to moderate reactions. Lastly, the "slow to warm up" profile encompassed low activity level, slow adaptability, low intensity of, and negativity in emotions. Note also, similar dimensions have been suggested by other temperament researchers such as Rothbart (1981): negative affectivity, effortful control and positive affectivity surgency; and Caspi (2000): under controlled children, well-adjusted children and inhibited children.

As can be seen in **Table 1**, the EFA results identified multiple cross-loads. Cross-loads were expected, as the scales included in the analysis did not comprise of a single item and often covered several domains of socio-emotional competence and hence were interlinked. In order to maintain conceptual integrity and ensure adequate coverage of each factor, the cross-loading variables were not dropped. The exception was the prosocial scale at the 2-year

DCW. While this variable loaded on the easy-going dimension slightly more (0.039) than the regulation dimension, since this subscale at 4.5-year DCW loaded exclusively on the regulation dimension, it was kept under the regulation dimension at 2-years to ensure consistency of the procedure (see **Table 2** for descriptive statistics of variables retained).

After grouping the variables, they were scaled. Variables that had negative connotation were multiplied by - 1, so that a higher score meant more adaptive emotions or behaviors. This was also necessary to allow later aggregation of the scales for the index (and is why Factor 1 in the EFA was ultimately labeled easy-going, despite the original factor loading structure in the EFA suggesting a more "difficult" child). After scaling and inverting (where necessary), the variables were averaged in order to calculate dimension scores of the index (easy-going, regulation and exuberance). Then the dimensions were aggregated additively, after standardization, to calculate the overall index at each DCW separately and also with all DCW taken together.

We then calculated two versions of dimension scores. In the first version, variables at each time point were weighted equally, while for the second version, weighting of the variables was

TABLE 3 | Characteristics of the sample before and after deletion and missing value analysis.

Characteristics	Full data (N = 6156)		Restricted dataset (N = 4839)	
	n	%	n	%
CHILD'S GENDER				
Boy	3165	51.4	2459	50.8
Girl	2986	48.5	2378	49.1
CHILD'S PARITY				
0	2170	35.3	1740	36.0
1–5	3328	54.1	2567	53.0
6–10	79	1.3	59	1.2
≥11	14	0.2	11	0.2
MOTHER'S AGE				
<20	250	4.1	199	4.1
20–24 years	829	13.5	657	13.6
25–29 years	1495	24.3	1131	23.4
30–34 years	1972	32.0	1587	32.8
35–39 years	1338	21.7	1058	21.9
≥40 years	271	4.4	207	4.3
MOTHER'S ETHNICITY				
European	3491	56.7	2927	60.5
Maori	813	13.2	622	12.9
Pacific	785	12.8	558	11.5
Asian	845	13.7	571	11.8
MELAA	115	1.9	79	1.6
Other	89	1.4	70	1.4
HOUSEHOLD DEPRIVATION QUINTILE INDEX 2013[a]				
NZ Dep Q1 (least deprived)	1227	19.9	1048	21.7
NZ Dep Q2	1119	18.2	948	19.6
NZ Dep Q3	1033	16.8	887	18.3
NZ Dep Q4	996	16.2	832	17.2
NZ Dep Q5 (most deprived)	1436	23.3	1095	22.6

[a]The New Zealand Deprivation Index combines nine socioeconomic variables from the 2006 census capturing eight dimensions of deprivation at the small area level. In this table the original deprivation scores (measured in deciles) have been collapsed into quintiles. Quintile 1 (Q1: Deciles 1 and 2) represents the households in the least deprived 20% areas, whereas Quintile 5 (Q5: Deciles 9 and 10) represents the households in the most deprived 20% areas (Salmond et al., 2007).

done according to variance estimates based on factor analysis (FA). Since the socio-emotional index at the previous two DCWs did not comprise variables based on observations, for the 4.5-year DCW, dimension scores were calculated with and without observation scores in order to see if observations complimented the scores based on parental ratings.

Next, factors calculated by aggregating variables weighted according to FA variance estimates were aggregated to give a socio-emotional competence score for each DCW, weighted according to the variance explained by that factor in comparison to the other(s). Hence, this process led to two versions of the socio-emotional competence index (one with equal and one with factor analysis based weighting) for the 9-month and 2-year DCWs and four index versions at 4.5-year DCW (as this included a version with and without observation data).

Socio-emotional competence at each DCW were further added together using two weighting methods to give overall index scores across all data collection waves: equal weighting and time-based weighting. For the later, the variables were weighted according to the distance of the DCW from 4.5-year DCW. This was done as it is both logical and intuitive to give scores at distant time-points less weight. Weights of 1, 1.33 $(1 + 15/45)$[1] and 2 $(1+45/45)$[2] were assigned to 9-month, 2-year and 4.5-year DCWs, respectively. Eight indices resulted from the analyses (see **Figure 2**):

1. Index with *equally* weighted socio-emotional competence scores at each DCW consisting of equally weighted dimension and variable scores *without* observations.
2. Index with *equally* weighted socio-emotional competence scores at each DCW consisting of equally weighted dimension and variable scores *with* observations.
3. Index with *equally* weighted socio-emotional competence scores at each DCW with *FA* variance estimates-based weighted dimension and variable scores *without* observations.
4. Index with *equally* weighted socio-emotional competence scores at each DCW with *FA* variance estimates-based weighted dimension and variable scores *with* observations.
5. Index with *time*-weighted socio-emotional competence scores at each DCW consisting of equally weighted dimension and variable scores *without* observations.
6. Index with *time*-weighted socio-emotional competence scores at each DCW consisting of equally weighted dimension and variable scores *with* observations.
7. Index with *time*-weighted socio-emotional competence scores at each DCW using *FA* variance estimates-based weighted dimensions and variable scores *without* observations.
8. Index with *time*-weighted socio-emotional competence scores at each DCW *with FA* variance estimates-based weighted dimension and variable scores *with* observation.

Testing of the Index

To briefly test each derived index and help select the most empirically sound index, we examined the relationship between these eight versions of the socio-emotional competence index with three other relevant outcome variables in the GUiNZ data set. These variables included: (1) handclap task at 4.5-year DCW (a measure of inhibitory control and executive memory functioning, which are aspects of cognitive functioning); (2) pragmatic language at 4.5-year DCW (measures the ability to communicate to others in more than just the vocabulary itself) and (3) parental perceived school readiness and difficulties experienced while starting school, which came from the 6-year DCW and allowed us to look at predictions over time. These outcome variables are described in **Appendix A**.

[1]9_month DCW was assigned a weight of 1. Since the gap between 9 months and 4.5 years is 45 months, and the gap between 9 months and 2 years is 15 months; 2-year DCW was assigned the weight of $1 + 15/45 = 1.33$, while 4.5-year DCW was assigned a weight of $1+45/45 = 2$.
[2]9_month DCW was assigned a weight of 1. Since the gap between 9 months and 4.5 years is 45 months, and the gap between 9 months and 2 years is 15 months; 2-year DCW was assigned the weight of $1 + 15/45 = 1.33$, while 4.5-year DCW was assigned a weight of $1+45/45 = 2$.

DCWs	Types of Weighting for Variables		Dimensions	Types of Weighting for Dimensions	Indices at Each DCW
9 months	Variables with Equal Weighting	→	Dimension 1 (Easy-Going)	Dimensions Weighed Equally	1. SEC Index (EW_9months)
			Dimension 2 (Regulation)		
	Variables with FA variance estimates-based Weighting	→	Dimension 1 (Easy-Going)	Dimensions Weighed Based on FA Variance	2. SEC Index (FA_9months)
			Dimension 2 (Regulation)		
2 years	Variables with Equal Weighting	→	Dimension 1(Easy-Going)	Dimensions Weighed Equally	3. SEC Index (EW_2years)
			Dimension 2 (Regulation)		
	Variables with FA variance estimates-based Weighting	→	Dimension 1(Easy-Going)	Dimensions Weighed Based on FA Variance	4. SEC Index (FA_2years)
			Dimension 2 (Regulation)		
4.5 years	Variables with Equal Weighting	→	Dimension 1(Easy-Going)	Dimensions Weighed Equally	5. SEC Index without observations (EW_4.5years)
			Dimension 2 (Regulation)		
			Dimension 3 (Exuberance) without observations		
			Dimension 1 (Easy-Going)	Dimensions Weighed Equally	6. SEC Index with Observations (EW_4.5years)
			Dimension 2 (Regulation)		
			Dimension 3 (Exuberance) with observations		
	Variables with FA variance estimates-based Weighting	→	Dimension 1(Easy-Going)	Dimensions Weighed Based on FA Variance	7. SEC Index without Observations (FA_4.5years)
			Dimension 2 (Regulation)		
			Dimension 3 (Exuberance) without observations		
			Dimension 1(Easy-Going)	Dimensions Weighed Based on FA Variance	8. SEC Index with Observations (FA_4.5years)
			Dimension 2 (Regulation)		
			Dimension 3 (Exuberance) with observations		

Equal Weighting	• Overall SEC Index including 1,3 & 5 (Index 1)	
	• Overall SEC Index including 1,3 & 6 (Index 2)	
	• Overall SEC Index including 2,4 & 7 (Index 3)	
	• Overall SEC Index including 2, 4 & 8 (Index 4)	
Time Weighting	• Overall SEC Index including 1,3 & 5 (Index 5)	
	• Overall SEC Index including 1,3 & 6 (Index 6)	
	• Overall SEC Index including 2,4 & 7 (Index 7)	
	• Overall SEC Index including 2, 4 & 8 (Index 8)	

FIGURE 2 | Showing approaches taken to calculate the index of socio-emotional competence. SEC, Socio-Emotional Competence; EW, Equal Weighting; FA, Factor Analysis.

We chose executive functioning because affective processes are purported to interact dynamically with attentive executive functioning and communication (Beauchamp and Anderson, 2010). For example, Schonert-Reichl et al. (2015) found that an intervention to improve social-emotional development in children resulted in improvements in cognitive control in children aged 9–11.16 years. With regard to language, Horwitz et al. (2003) found language delays in children aged 12–39 months were associated with poor social competence. Similarly, Irwin et al. (2002) reported poor social-emotional adjustment in late-talking toddlers. When interventions were carried out to improve early language, a reduction in internalizing, externalizing and overall problematic behaviors of preschool children was seen one 1 year after the intervention (Curtis et al., 2017).

As for school readiness, pre-literacy skills such as phonological awareness, and alphabet and print knowledge, which many argue signify academic preparedness, have been found to be predicted by emotional expressiveness, regulation, and knowledge in preschool children (Curby et al., 2015). Denham et al. (2015) found that emotional knowledge enhanced early classroom adjustment via social-competence, while Izard et al. (2001) reported that children from economically disadvantaged backgrounds tended to do better academically if they had adequate emotional knowledge.

As expected, all 8 indices were positively associated with the all three outcome variables (see **Table 4**), but index 6 with time-weighted socio-emotional competence scores (at each DCW), and equally weighted dimensions and variable scores including observations, were found to be the most empirically sound (with relatively better associations). As per the guidelines of Evans (1996), Index 6 had weak to moderate levels of correlation with the three outcome variables. The strongest correlation was with pragmatic language.

There are a number of reasons as to why Index 6 emerged as the most parsimonious one. It included variables related to observations such as expression identification on the AKT measure and observations related to child's mood and engagement with the assessor (at 4.5-year DCW). Observations are considered to add breadth to the assessment, and are advised to be used in conjunction with parental reports (Carter et al., 2003, 2004; Rothbart, 2011). In addition, as mentioned above, Index 6 also comprised of time-weighted socio-emotional competence scores at each DCW, with variables and the dimensions weighted equally as opposed to using empirical weighting. Equal weighting is generally the preferred mode of weighting variables as it gives better predictive results (Babbie, 1995; Booysen, 2002; Sanson et al., 2005; Blakemore and Gibbings, 2006; Misson et al., 2011). In addition, it also makes intuitive sense for the time weighted overall index to be statistically the most sound, as relatively less weight should

TABLE 4 | Correlations between the eight different indices and handclap total score (4.5 years), pragmatic language (4.5 years), and school readiness (6 years).

Indices	Hand clap task (4.5 years; $n = 4668$)	Pragmatic language (4.5 years; $n = 4837$)	School readiness (6 years; $n = 4464$)
1. Overall socio-emotional competence equal weighting	0.118***	0.336***	0.194***
2. Overall socio-emotional competence equal weighting with observations	0.143***	0.368***	0.210***
3. Overall socio-emotional competence equal weighting (FA based)	0.131***	0.289***	0.158***
4. Overall socio-emotional competence equal weighting with observations (FA based)	0.149***	0.310***	0.169***
5. Overall socio-emotional competence time weighting	0.157***	0.363***	0.195***
6. Overall socio-emotional competence time weighting with observations	0.195***	0.402***	0.209***
7. Overall socio-emotional competence time weighting (FA based)	0.169***	0.355***	0.194***
8. Overall socio-emotional competence time weighting with observations (FA based)	0.179***	0.366***	0.199***

*** $p < 0.001$.

be given to socio-emotional competence scores at distant time points especially in a fast changing and developing child.

The socio-emotional competence Index 6 (refer to **Table 4**) was then transformed to have a mean of 100 and standard deviation of 15. Following which, the index score at and across the three DCWs was categorized by using 1 SD as the cut off to form three categories of socio-emotional competence: Low, Average, and High. This categorical data can be used to calculate trajectories of change or stability in levels of socio-emotional competence across the three DCWs, further allowing later exploration of demographical, familial, and other environmental factors that contribute toward these patterns of stability or change. In addition, these trajectories can be used to estimate their influence on other related developmental outcomes.

Below, we briefly explore the relationships between the index and pragmatic language and maternal perceived school readiness. We do not explore the relationship between the index, and executive functioning as this is explored in greater detail in a later study (Ahmad et al., 2018a) and hence, is not detailed here.

The Socio-Emotional Competence Index, Pragmatic Language and School Readiness

Overall, we found that socio-emotional competence at each DCW related to pragmatic language and school readiness (maternal perceived) and that this association grew stronger with time (see **Table 5**). Also, relatively stronger correlations were observed between socio-emotional competence scores at 2 years and 4.5 years compared to their relationship with socio-emotional competence score at 9 months (see **Table 6**). This probably reflects the fact that the first year of a child's life is marked by rapid changes in their development and their ability to express themselves. These changes tend to become relatively less rapid beyond infancy (Carter et al., 2003, 2004). In addition,

TABLE 5 | Pearson correlations between dimensions (Easy-going, Regulation and Exuberance) and overall scores of socio-emotional competence scores at each DCW with pragmatic language (4.5 years) and school readiness (6 years).

Dimensions	Pragmatic language (4.5 years; $n = 4837$)	School readiness (6 years; $N = 4402$)
9 MONTHS		
Easy-going	0.08***	0.05***
Regulation	0.13***	0.10***
Socio-emotional competence score	0.15***	0.10***
2 YEARS		
Easy-going	0.28***	0.13***
Regulation	0.14***	0.06***
Socio-emotional competence score	0.24***	0.11***
4.5 YEARS		
Easy-going	0.25***	0.16***
Regulation	0.33***	0.17***
Exuberance	0.13***	0.06***
Socio-emotional competence score	0.42***	0.21***

*** $p < 0.001$.

socio-emotional competence at the 9-month DCW was measured by using scales measuring different aspects of temperament, which is only moderately stable in the first 2 years of a child's life (Bates and Pettit, 2007; Rothbart, 2011); thereby, explaining stronger relationships of socio-emotional competence scores at the later DCWs with each other, and the outcome variables.

With respect to the three socio-emotional dimensions correlating with our outcome measures, relatively stronger correlations were found with pragmatic language compared to school readiness (see **Table 5**). The strongest relationship was

TABLE 6 | Showing pearson correlations between socio-emotional competence index score at each DCW (N = 4839).

Socio-emotional competence index score at each DCW	1	2
9 months		
2 years	0.27***	
4.5 years	0.27***	0.41***

***p < 0.001.

TABLE 7 | Showing pearson correlations between socio-emotional competence dimensions (Easy-going, regulation and exuberance) at Each DCW (N = 4839).

Child dimensions	1	2	3	4	5	6
9 MONTHS						
Easy-going						
Regulation	−0.02					
2 YEARS						
Easy-going	0.30***	0.06***				
Regulation	−0.02	0.24***	0.16***			
4.5 YEARS						
Easy-going	0.26***	0.01	0.53***	0.01		
Regulation	0.03***	0.26***	0.18***	0.33***	0.16***	
Exuberance	0.14***	0.04***	0.17***	0.01	0.15***	0.04***

***p < 0.001.

observed for pragmatic language with the regulation dimension at the 4.5-year DCW. This is perhaps not surprising given that even in early development, words are considered to be important tools for self-regulation as they are believed to be a means through which children can positively express their needs and dissatisfaction or frustration, rather than doing something impulsive (Cournoyer et al., 1998; Vallotton and Ayoub, 2011). The second highest correlation was observed between the easy-going dimension at 2 years and pragmatic language at 4.5 years (see **Table 5**). The easy-going dimension at the 2 year DCW included a vocabulary variable (words related to emotions), which may have enhanced the association; however, a similarly strong relationship was observed between pragmatic language and the easy-going dimension at 4.5 years, which did not contain a language related variable in the index. This suggests that the association between an easy-going dimension and pragmatic language is not just a method effect and that easy-going children may be better at communicating. There is some support for the association between the easy-going dimension and pragmatic language in the literature. Children with pragmatic language competence can typically use language in context, and the ability to do this has been strongly associated with decreased behavioral problems in a community sample (Ketelaars et al., 2010; Rints et al., 2015).

The relationships between socio-emotional competence dimensions at each DCW are shown in **Table 7**. Association were generally stronger between the same dimensions at each DCW. The dimensions that did not significantly overlap all involved either the 9 month or 2 year regulation dimension. The lack of statistically significant associations could be because the first few years of a child's life are marked by rapid development and this development may not be uniform across all the socio-emotional domains, perhaps particularly regulation (Sanson et al., 2010). The lack of associations with regulation may also be because what is generally considered dysregulated behavior (such as tantrums, non-compliance and aggression toward peers), and therefore maladaptive, may in fact be a hallmark of normal development during the early years, reflecting in particular the development of independence around the age of two (Dunn et al., 1996).

Overall, the correlations between the dimensions were not high over time and this may be because parental reports were exclusively relied on to assess socio-emotional development at the 9-month and 2-year DCW. Parents may not have the

experience to discern between problematic and non-problematic dysregulation (Carter et al., 2003) as they interpret their child's behavior based on their history of interaction with children, as well as cultural values and beliefs (Clark et al., 2004; Rothbart, 2011). Notably, at the 4.5-year DCW, no correlation was observed between the regulation and the exuberance dimensions. While this could be because these dimensions measure different aspects of socio-emotional competence, it may also be a method effect as the exuberance dimension included observations, which may differ from parental ratings (Carter et al., 2003), which were exclusively relied upon to calculate the regulation child dimension.

Strengths and Limitations

The strengths of this index are that the measures used to construct the index are easily accessible and unproblematic to administer to different populations and in different settings making it possible to repeat this measure to look for change if required. In addition, different versions of the index (using different types of weighting) were calculated and assessed, which adds to the validity of the index. The final index chosen is sensitive to different levels of socio-emotional competence, as at least 15.3% of the sample fell into either low or high categories of socio-emotional competence across all the DCWs and this allows for potential tracking of socio-emotional competence over time. Finally, the index uses measures that tap into emotional and social strengths as well-maladaptive behaviors, allowing for a more comprehensive assessment, which may, over time be helpful to those looking to identify children at potential risk of poor socio-emotional development.

In our study, a potential limitation is that the index was constructed mostly using maternal ratings, apart from the observations included in the 4.5-year DCW. Paternal ratings were not included. This singular reliance on parental report is not ideal (Carter et al., 2003) as there may be biases involved in their ratings (Meisels, 1998; Campbell et al., 2016). However, the measures used were mostly focused on factual information regarding specific situations, which decreased the reliance on parental interpretation (Campbell et al., 2016).

Furthermore, parents' views about their children affect their parental practices, ultimately influencing children's behaviors and vice versa, potentially resulting in a feedback loop (Sameroff and Chandler, 1975; Thomas and Chess, 1977; Sameroff, 2009). Therefore, parental reports are important to capture when trying to create a picture of a child's socio-emotional development.

Another drawback was that, for developmental reasons, different variables were used to measure emotions and behaviors at each DCW. Therefore, the meaning of dimensions and the socio-emotional index at each DCW may differ from one and other. However, as mentioned before, consistently stronger relationships were found between the same dimensions at each DCW (**Table 7**), which indicates they are likely to be measuring a similar, if not the same, basic concept. It is important to keep in mind that the cut-off of 1SD used to categorize socio-emotional development in to high and low categories was arbitrary, which means it cannot be considered clinically significant. Therefore, although it still provides an avenue to compare children's socio-emotional competence scores relative to one another, researchers need to treat such data sensitively.

At a more general level, it is important to acknowledge that there is continued debate about the value of index formation as it is argued to simplify a complex concept and thereby lose information. Despite this, we believe we have created a potentially useful index, but how important a child's score on this index is in the long term is yet to be determined. Caution should be exercised when interpreting the results and using the index. It is important to remember that early childhood is a dynamic phase of a child's life. Changes are taking place in the child's mind and body at a greater rate than any other life-stage. Therefore, while our index can provide a snapshot of a pre-school child's life at a particular moment, it is important to remember their socio-emotional strengths are likely to change. Secondly, we need to be mindful that the variables included in the index assess emotions and behaviors within a largely Western framework. That is, our variables reward emotions and behaviors considered generally "adaptive" in the Western world, which may be viewed differently under different cultural frameworks. Therefore, we believe that while this index provides a potentially useful early measure of "socio-emotional competence" of a child in this cohort, researchers and policy makers need to be aware that changes may be needed when applying the index to other contexts.

CONCLUSION

An Index of socio-emotional competence was developed using data from the GUiNZ study for children aged approximately 9 months to 4.5 years. The Index was calculated following the methods described by Booysen (2002) and the Joint Research Centre-European Commission (2008). This paper initially aimed to develop an index of socio-emotional development based on the model of socio-emotional competence given by Denham (2006), which drew on the Rose-Krasnor (1997) model of social competence, comprising of emotional expressiveness, understanding of emotions, regulation of emotions and behavior, social problem solving and social relationship skills. However, factor analysis led to the emergence of three child dimensions: easy-going, regulation and exuberance (the latter emerged at the 4.5-year DCW only as a child's emotions and behaviors become more complex with development). The emergence of dimensions rather than domains/components of socio-emotional competence (given by Denham, 2006) was in part expected because the variables used to construct the index each concerned more than one aspect of socio-emotional competence. The variables included provided good conceptual coverage of the socio-emotional competence model, and had adequate internal consistency and discriminatory power. However, variables at each DCW were not the same, reflecting the development of the child at each DCW.

The index not only gives an overall score of socio-emotional competence taking into account data from each DCW but also gives socio-emotional competence score at each DCW. It also gives a score of each of the socio-emotional competence dimensions (easy-going, regulation and exuberance), which allows for analysis of weakness or strength in each dimension—adding breadth to the data. We also calculated both continuous and categorical index scores for socio-emotional competence at each DCW and across all DCWs, which allow for flexibility in further analytical use of the index.

Our socio-emotional competence index at each DCW also gives us access to contemporary and localized information about the relative prevalence, stability and change in socio-emotional competence in New Zealand children across the preschool years (this is explored in another article by Ahmad et al., 2018b). Future research can then explore the demographical, familial and environmental basis of the trajectory of change or stability in socio-emotional competence. The knowledge gained from this research can potentially be used to identify children at risk of poor socio-emotional competence development as early as possible and to explore which factors may help to mitigate persistently low levels of socio-emotional competence and which factors promote improvements. In addition, the index could be used to provide a way of potentially tracking change in socio-emotional competence over the preschool years if an intervention (whether it be targeted or universal) was undertaken.

This index can potentially also be used to assess how different patterns of early socio-emotional development affect other outcomes. Ultimately, we hope that the use of this index will help us to better understand the development of children's socio-emotional competence over time, with a view to later helping children, families, and communities to reap the lifelong benefits that socio-emotional competence can bring.

ETHICS STATEMENT

Each data collection wave of GUiNZ was initiated after getting permission from the New Zealand Ministry of Health Northern

Y Regional Ethics Committee (NTY/08/06/055). The study involves anonymized datasets. Written informed consent was taken from the participating parents, while they also provided written informed consent on the behalf of their child who participated in the study. Participants had full rights to withdraw from the research at any time without any penalty.

AUTHOR CONTRIBUTIONS

REFERENCES

Achenbach, T. M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H. C., et al. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: research findings, applications, and future directions. *J. Child Psychol. Psychiatry* 49, 251–275. doi: 10.1111/j.1469-7610.2007.01867.x

Ahmad, S., Peterson, E. R., Waldie, K. E., and Morton, S. M. B. (2018a). *Prevalence and Trajectories of Change/Stability in Socio-emotional Competence for Preschoolers.* Unpublished manuscript, School of Psychology; Faculty of Science; The University of Auckland, Auckland.

Ahmad, S., Peterson, E. R., Waldie, K. E., and Morton, S. M. B. (2018b). *Do Trajectories of Change and Stability in Socio-emotional Competence During Early Childhood Predict Executive Function at 4.5 Years?* Unpublished manuscript, School of Psychology; Faculty of Science; The University of Auckland, Auckland.

Babbie, E. (1995). *The Practice of Social Research.* Washington, DC: Wadsworth.

Baron-Cohen, S., Baron-Cohen, S., Ring, H. A., Ring, H. A., Wheelright, S., Wheelright, S., et al. (1999). Social intelligence in the normal and autistic brain: an fMRI study. *Eur. J. Neurosci.* 11, 1891–1898. doi: 10.1046/j.1460-9568.1999.00621.x

Bates, J. E., and Pettit, G. S. (2007). "Temperament, parenting, and socialization," in *Handbook of Socialization: Theory and Research*, eds J. E. Grusec and P. D. Hastings (New York, NY: Guilford Press), 153–177.

Beauchamp, M. H., and Anderson, V. (2010). SOCIAL: an integrative framework for the development of social skills. *Psychol. Bull.* 136, 39–64. doi: 10.1037/a0017768

Blakemore, T., and Gibbings, J. (2006). *Measuring the Socio-Economic Position of Families in HILDA and LSAC.* ACSPRI Social Science Methodology Conference, University of Sydney. Available online at: https://old.acspri.org.au/conference2006/proceedings/streams/SEP%20paper.%2013_11_06.pdf

Booysen, F. (2002). An overview and evaluation of composite indices of development. *Soc. Ind. Res.* 59, 115–151. doi: 10.1023/A:1016275505152

Britto, P. R. (2012). *School Readiness: A Conceptual Network (UNICEF).* Available online at: http://www.unicef.org/education/files/Chil2Child_ConceptualFramework_FINAL%281%29.pdf

Brown, J. D. (2011). Likert items and scales of measurement? *SHIKEN JALT Test. Eval. SIG Newsl.* 15, 10–14. doi: 10.5032/jae.1994.04031

Calaguas, G. M., and Dizon, C. S. (2011). Development and initial validation of the Social Competency Inventory for tertiary level faculty members. *Int. J. Hum. Soc. Sci.* 6, 171–176.

Campbell, S. B., Denham, S. A., Howarth, G. Z., Jones, S. M., Whittaker, J. V., Williford, A. P., et al. (2016). Commentary on the review of measures of early childhood social and emotional development: conceptualization, critique, and recommendations. *J. Appl. Dev. Psychol.* 45, 19–41. doi: 10.1016/j.appdev.2016.01.008

Carter, A.S., Briggs-Gowan, M.J., Jones, S.M., and Little, T. D. (2003). The infant-toddler social and emotional assessment: factor structure, reliability, and validity. *J. Abnorm. Child Psychol.* 31, 495–514. doi: 10.1023/A:1025449031360

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. *J. Child Psychol. Psychiatry* 45, 109–134. doi: 10.1046/j.0021-9630.2003.00316.x

Caspi, A. (2000). The child is the father of man: personality continuities from childhood to adulthood. *J. Personal. Soc. Psychol.* 78, 158–172. doi: 10.1037/0022-3514.78.1.158

Clark, R., Tluczek, A., and Gallagher, K. (2004). "Assessment of parent-child early relational disturbances," in *Handbook of Infant, Toddler, and Preschool Mental Health,* eds R. Del Carmen-Wiggens and A. Carter (New York, NY: Oxford University Press), 25–60.

Clarke-Stewart, A., and Parke, R. S. (2014). *Social Development.* Hoboken, NJ: Wiley.

Cohen, J. (1988). *Statistical Power Analysis for the Behaviorial Sciences, 2nd Edn.* Hillsdale, NJ: L. Erbaum.

Cohen, J., Onunaku, N., Clothier, S., and Poppe, J. (2005). *Helping Young Children Succeed: Strategies to Promote Early Childhood Social and Emotional Development,* Washington, DC: National Conference of State Legislatures.

Costello, A. B., and Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval.* 10, 1–9.

Cournoyer, M., Solomon, C. R., and Trudel, M. (1998). "Je parle donc j'attends?": Language et autocontrôle ches le jeune enfant. [I speak then I expect: Language and self-control in the young child at home]. *Can. J. Behav. Sci.* 30, 69–81.

Cowan, C. R., Hauser, R. M., Kominski, R. A., Levin, H. M., Lucas, S. R., Morgan, S. L., et al. (2012). *Improving the Measurement of Socioeconomic Status for the National Assessment of Educational Progress: A theoretical foundation.* Washington, DC: National Center for Education Statistics.

Croft, S., Stride, C., Maughan, B., and Rowe, R. (2015). Validity of the strengths and difficulties questionnaire in preschool-aged children. *Pediatrics* 135, 1210–1219. doi: 10.1542/peds.2014-2920

Curby, T. W., Brown, C. A., Bassett, H. H., and Denham, S. A. (2015). Associations between preschooler's social-emotional competence and preliteracy skills. *Infant Child Dev.* 24, 549–570. doi: 10.1002/icd.1899

Curtis, P. R., Kaiser, A. P., Estabrook, R., and Roberts, M. Y. (2017). The longitudinal effects of early language intervention on children's problem behaviors. *Child Dev.* doi: 10.1111/cdev.12942. [Epub ahead of print].

de la Osa, N., Granero, R., Penelo, E., Domenech, J. M., and Ezpeleta, L. (2013). The short and very short forms of the children's behavior questionnaire in a community sample of preschoolers. *Assessment* 21, 436–476. doi: 10.1177/1073191113508809

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the {EM} algorithm. *J. R. Stat. Soc.* 39, 1–38. doi: 10.1111/j.2517-6161.1977.tb01600.x

Denham, S.A. (1998). *Emotional Development in Young Children.* New York, NY: Guilford

Denham, S. A. (1986). Social cognition, social behavior, and emotion in pre-schoolers:contextual validation. *Child Dev.* 57, 194–201. doi: 10.2307/1130651

Denham, S. A. (2006). Social relationships and school readiness. *Early Edu. Dev.* 17, 151–176. doi: 10.1207/s15566935eed1701

Denham, S. A., Bassett, H. H., Brown, C., Way, E., and Steed, J. (2015). "I Know How You Feel": preschoolers' emotion knowledge contributes to early school success. *J. Early Childh. Res.* 13, 252–262. doi: 10.1177/1476718X13497354

Denham, S. A., Bassett, H. H., Way, E., Mincic, M., Zinsser, K., and Graling, K. (2012). Preschoolers' emotion knowledge: self-regulatory foundations, and predictions of early school success. *Cogn. Emot.* 26, 667–679. doi: 10.1080/02699931.2011.602049

Denham, S. A., Cook, M., and Zoller, D. (1992). "Baby looks very sad": implications of conversations about feelings between mother and preschooler. *Br. J. Dev. Psychol.* 10, 301–315. doi: 10.1111/j.2044-835X.1992.tb00579.x

DesRosiers, F.S., and Busch-Rossnagel, N. A. (1997). Self concept in toddlers. *Infant Young Child.* 10, 15–26. doi: 10.1097/00001163-199707000-00004

Diamond, A., and Taylor, C. (1996). Development of an aspect of executive control: development of the abilities to remember what I said and to "Do as I say, not as I do". *Dev. Psychobiol.* 29, 315–334.

Dickinson, D. K., McCabe, A., and Sprague, K. (2001). *Teacher Rating of Oral Language and Literacy (TROLL): A Research-Based Tool*. Ciera Report #3-016. Michigan, US: Centre for the Improvement of Early Reading Achievement (CIERA), University of Michigan. Available online at: www.ciera.org/library/reports/inquiry-3/3-016.pdf

Diener, E. D., and Suh, E. (1997). Measuring quality of life: economic, social, and subjective indicators. *Soc. Indicat. Res.* 40, 189–216. doi: 10.1023/A:1006859511756

Drewnowski, J. (1974). "On measuring and planning the quality," in *Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies*, eds N. Wenger (New York: Le Jacq), 170–83

D'Souza, S., Waldie, K. E., Peterson, E. R., Underwood, L., and Morton, S. M. B. (2017). Psychometric properties and normative data for the preschool strengths and difficulties questionnaire in two-year-old children. *J. Abnorm. Child Psychol.* 45, 345–357. doi: 10.1007/s10802-016-0176-2

Dunn, J., Creps, C., and Brown, J. (1996). Children's family relationships between two and five: developmental changes and individual differences. *Soc. Dev.* 5, 230–250.

Evans, J. D. (1996). *Straightforward Statistics for the Behavioral Sciences*. Pacific Grove, CA: Brooks/Cole Publishing.

Fan, X., and Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivar. Behav. Res.* 42, 509–529. doi: 10.1080/00273170701382864

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., and Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Appl. Psycholinguist.* 21, 95–115. doi: 10.1017/S0142716400001053

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics: and Sex and Drugs and Rock 'n' roll, 4th Edn.* London: Sage.

Golden, C. , J., Hammeke, T. A., and Purisch, A. D. (1979). *The Standardized Luria-Nebraska Neuropsychological Battery: A manual for clinical and experimental use.* Lincoln: University of Nebraska Press.

Goodman, A., Lamping, D. L., and Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesized five subscales on the strengths and difficulties questionnaire (SDQ): data from British parents, teachers and children. *J. Abnorm. Child Psychol.* 38, 1179–1191. doi: 10.1007/s10802-010-9434-x

Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *J. Child Psychol. Psychiatry* 38, 581–586. doi: 10.1111/j.1469-7610.1997.tb01545.x

Halberstadt, A. G., Denham, S. A., and Dunsmore, J. C. (2001). Affective social competence. *Soc. Dev.* 10, 79–119. doi: 10.1111/1467-9507.00150

Han, H. S., and Kemple, K. M. (2006). Components of social competence and strategies of support: considering what to teach and how. *Early Childhood Educ. J.* 34, 241–246. doi: 10.1007/s10643-006-0139-2

Harris, P. (1989). *Children and Emotion*. Oxford: Basil Blackwell.

He, J. P., Burstein, M., Schmitz, A., and Merikangas, K. R. (2013). The strengths and difficulties questionnaire (SDQ): the factor structure and scale validation in U.S. *Adoles. J. Abnorm. Child Psychol.* 41, 583–595. doi: 10.1007/s10802-012-9696-6

Hooper, D., Coughlan, J., and Mullen, M. (2008). Structural equation modelling : guidelines for determining model fit structural equation modelling : guidelines for determining model fit. *Electron. J. Bus. Res. Methods* 6, 53–60. doi: 10.1037/1082-989X.12.1.58

Horwitz, S. M., Irwin, J. R., Briggs-Gowan, M. J., Bosson Heenan, J. M., Mendoza, J., and Carter, A. S. (2003). Language delay in a community cohort of young children. *J. Am. Acad. Child Adol. Psychiatry* 42, 932–940. doi: 10.1097/01.CHI.0000046889.27264.5E

Hollingshead, A. B. (1975). *Four Factor Index of Social Status*. New Haven, CT: University Press.

Hughes, S.O., Shewchuk, R. M., Baskin, M.L., Nicklas, T. A., and Haiyan, Q. (2008). Indulgent feeding style and children's weight status in preschool. *J. Dev. Behav. Pediatrics* 29, 403–410. doi: 10.1097/DBP.0b013e318182a976

Irwin, J. R., Carter, A. S., and Briggs-Gowan, M. J. (2002). The social-emotional development of late-talking toddlers. *J. Am. Acad. Child Adol. Psychiatry* 41, 1324–1332. doi: 10.1097/00004583-200211000-00014

Izard, C. E., Fine, S., Schultz, D., Mostow, A. J., Ackerman, B., and Youngstrom, E. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychol. Sci.* 12, 18–23. doi: 10.1111/1467-9280.00304

Jamieson, S. (2004). Likert scales: how to (ab)use them. *Med. Edu.* 38, 1217–1218. doi: 10.1111/j.1365-2929.2004.02012.x

Joint Research Centre-European Commission (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide.* Paris: OECD Publishing.

Kestenbaum, R., and Gelman, S. A. (1995). Preschool children's identification and understanding of mixed emotions. *Cogn. Dev.* 10, 443–458. doi:10.1016/0885-2014(95)90006-3

Ketelaars, M. P., Cuperus, J., Jansonius, K., and Verhoeven, L. (2010). Pragmatic language impairment and associated behavioural problems. *Int. J. Lang. Commun. Disord.* 45, 204–214. *doi*: 10.3109/13682820902863090

Knapp, T. R. (1990). Treating ordinal scales as interval scales: an attempt to resolve the controversy. *Nurs. Res.* 39, 121–123. doi: 10.1097/00006199-199003000-00019

Kopp, C. B. (1989). Regulation of distress and negative emotions: a development view. *Dev. Psychol.* 25, 343–354. doi: 10.1037//0012-1649.25.3.343

Land, K. C., Lamb, V. L., Meadows, S. O., and Taylor, A. (2007). Measuring trends in child well-being: an evidence-based approach. *Soc. Ind. Res.* 80, 105–132. doi: 10.1007/s11205-006-9023-0

Leffert, N., Benson, P. L., and Roehlkepartain, J. L. (1997). *Starting Out right: Developmental Assets for Children*. Minneapolis, MN: Search Institute.

McGranahan, D. V., Richard-Proust, C., Sovani, N. V., and Subramanian, M. (1972). *Contents and Measurement of Socio-Economic Development.* New York, NY: Praeger for UNRISD

Misson, S., Sanson, A., Berthelsen, D., Rogers, H., Rothman, S., Sipthorp, M., et al. (2011). *Tracking children's Development Over Time: The Longitudinal Study of Australian Children Outcome Indices, Waves 2 and 3.* Melbourne, VIC: Australian Institute of Family Studies.

Meisels, S. J. (1998). *Assessing Readiness.* (Report No. 3-002). Ann Arbor, MI: Center for the Improvement of Early Reading Achievement. Available online at: http://www.ciera.org/products/meisels-1998/reports32.html

Morgan, J. K., Izard, C.E., and King, K.A. (2009). Construct validity of the emotion matching task: preliminary evidence for convergent and criterion validity of a new emotion knowledge measure for young children. *Soc. Dev,* 19, 52–70. doi: 10.1111/j.1467-9507.2008.00529.x.

Morton, S. M.., Grant, C. C., Berry, S. D., Walker, C., Corkin, M., Ly, K., et al. (2017). *Growing Up in New Zealand: A Longitudinal Study of New Zealand Children and Their Families. Now We Are Four: Describing the Preschool Years.* Auckland: Growing Up in New Zealand. Available online at: https://cdn.auckland.ac.nz/assets/growingup/research-findings-impact/GUiNZ_Now%20we%20are%20four%20report.pdf

Morton, S. M. B., Atatoa Carr, P. E., Grant, C. C., Berry, S. D., Bandara, D. K., Mohal, J., et al. (2014). *Now We Are Two: Describing Our First 1000 days.* Auckland: Growing Up in New Zealand. Available online at: https://cdn.auckland.ac.nz/assets/growingup/research-findings-impact/report03.pdf

Niclasen, J., Skovgaard, A. M., Nybo-Andersen, A., Somhovd, M. J., and Obel, C. (2013). A confirmatory approach to examining the factor structure of the Strengths and Difficulties Questionnaire (SDQ): a large scale cohort study. *J. Abnorm. Child Psychol.* 41, 355–365. doi: 10.1007/s10802-012-9683-y

Nicoletti, G., Scarpetta, S., and Boylaud, O. (2000). *Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation. OECD.* Economics department working papers No. 226. Available online at: https://www.oecd.org/eco/outlook/1880867.pdf doi: 10.2139/ssrn.201668

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Adv. Health Sci. Edu.* 15, 625–632. doi: 10.1007/s10459-010-9222-y

Nunnally, J. C., Bernstein, I. H., and Begre, J. M. (1967). *Psychometric Theory, Vol. 226.* New York, NY: McGraw-Hill. Available online at: http://rds.epi-ucsf.org/ticr/syllabus/courses/46/2005/10/20/Lecture/readings/Psychometric%20Theory.pdf.

Peterson, E. R., Mohal, J., Waldie, K. E., Reese, E., Atatoa Carr, P. E., Grant, C. C., et al. (2017a). A cross-cultural analysis of the infant behavior questionnaire very short form: an item response theory analysis of infant temperament in new zealand. *J. Personal. Assessment.* 99, 574–584. doi: 10.1080/00223891.2017.1288128

Peterson, E. R., Waldie, K. E., Mohal, J., Reese, E., Atatoa Carr, P. E., Grant, C. C., et al. (2017b). Infant behavior questionnaire–revised very short form: a new factor structure's associations with parenting

perceptions and child language outcomes. *J. Personal. Assessment* 99, 561–573. doi: 10.1080/00223891.2017.1287709

Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., and Leerkes, E. (2014). Development and assessment of short and very short forms of the infant behavior questionnaire-revised. *J. Personal. Assess.* 96, 445–458. doi: 10.1080/00223891.2013.841171

Putnam, S. P., and Rothbart, M. K. (2006). Development of short and very short forms of the children's behavior questionnaire. *J. Personal. Assessment.* 87, 103–113. doi: 10.1207/s15327752jpa8701

Ram, R. (1982). Composite indices of physical quality of life, basic needs fulfillment, and income. A "principal component" representation. *J. Dev. Econo.* 11, 227–247. doi: 10.1016/0304-3878(82)90005-0

Rints, A., McAuley, T., and Nilsen, E. S. (2015). Social communication is predicted by inhibitory ability and adhd traits in preschool-aged children: a mediation model. *J. Attention Disord.* 19, 901–911. *doi*: 10.1177/1087054714558873

Roid, G. H., and Miller, L. J. (1997). *Social Emotional Rating Scale – Examiner Version. Leiter International Performance Scale – Revised (Leiter-R).* Wood Dale, IL: Stoelting.

Rose-Krasnor, L. (1997). The nature of social competence: a theoretical review. *Soc. Dev.* 6, 112–135. doi: 10.1111/j.1467-9507.1997.tb00097.x

Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Dev.* 52, 569–578. doi: 10.2307/1129176

Rothbart, M. K. (2011) *Becoming Who We Are: Temperament and Personality in Develoment.* New York, NY: United States of America: The Guilford.

Sameroff, A., and Chandler, M. (1975). "Reproductive risk and the continuum of caretaking casualty," in *Review of Child Development Research*, Vol. 4, eds F. Horowitz, (Chicago, IL: University of Chicago Press), 187–244.

Sameroff, A. J. (eds.). (2009). *The Transactional Model of Development: How Children And Contexts Shape Each Other.* Wash ington, DC: American Psychological Association. doi: 10.1037/11877-000

Salmond, C., Crampton, P., and Atkinson, J. (2007). *NZDep2006 Index of Deprivation.* Wellington: Department of Public Health, University of Otago.

Sanson, A., Misson, S., and the LSAC Outcome Index Working Group. (2005). *LSAC Technical Paper No. 2: Summarizing children's wellbeing. The LSAC Outcome Index.* Melbourne, VIC: Australian Institute of Family Studies.

Sanson, A. V., Misson, S., Hawkins, M. T., and Berthelsen, D. (2010). The development and validation of Australian indices of child development - Part I: conceptualisation and development. *Child Ind. Res. 3*, 275–292. doi:10.1007/s12187-009-9058-2

Schafer, J. L., and Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behav. Res.* 33, 545–571. doi: 10.1207/s15327906mbr3304_5

Schonert-Reichl, K. A., Oberle, E., Lawlor, M. S., Abbott, D., Thomson, K., Oberlander, T.F., et al. (2015). Enhancing cognitive and social-emotional development through a simple-to-administer mindfulness-based school program for elementary school children: a randomized controlled trial. *Dev. Psychol.* 51, 52–66. doi: 10.1037/a0038454

Shatz, M. (1994). *A Toddler's Life: Becoming a Person.* Oxford: Oxford University Press.

Slottje, D. J. (1991). Measuring the quality of life across countries. *Rev. Econ. Stat.* 73, 684–693. doi: 10.2307/2109407

Smith-Donald, R., Raver, C. C., Hayes, T., and Richardson, B. (2007). Preliminary construct and concurrent validity of the preschool self-regulation

assessment (PSRA) for field-based research. *Early Childh. Res. Q.* 22, 173–187. doi: 10.1016/j.ecresq.2007.01.002

Stubbing, J., Peterson, E. R., Waldie, K. E., Morton, S. M. B. (2017). *Temperament Development From 9 Months to 4.5 Years Using the CHILD Behaviour Questionnaire.* Unpublished manuscript, School of Psychology; Faculty of Science; The University of Auckland, Auckland.

Stevens, J. P. (2002). *Applied Multivariate Statistics For the Social Sciences, 4th edn.* Hillsdale, NS: Erlbaum.

Sullivan, G. M., and Artino, A. R. (2013). Analyzing and interpreting data from Likert-type scales. *J. Grad. Med. Edu.* 5, 541–542. doi: 10.4300/JGME-5-4-18

Tabachnick, B. G., and Fidell, L. S. (2001). *Using Multivariate Statistics. 4th edn.* Needham Heights, MA: Allyn and Bacon.

Tanaka, J. S. (1987). "How big is big enough?" sample size and goodness of fit in structural equaltion Models with latent variables. *Child Dev.* 58, 134–146.

Thomas, A., and Chess, S. (1977).*Temperament and Development.* New York, NY: Brunner/Mazel.

Thomas, A., Chess, S., and Birch, H. G. (1970). The origins of personality. *Sci. Am.* 223, 102–109.

Ul Haq, M. (1995). *Reflections on Human Development.* New York, NY: Oxford University Press.

Vallotton, C., and Ayoub, C. (2011). Use your words: The role of language in the development of toddlers' self-regulation. *Early Childh. Res. Q.* 26, 169–181. doi: 10.1016/j.ecresq.2010.09.002

Wakschlag, L. S., Leventhal, B. L., Briggs-Gowan, M. J., Danis, B., Keenan, K., Hill, C., et al. (2005). Defining the "disruptive" in preschool behavior: what diagnostic observation can teach us. *Clin. Child Fam. Psychol. Rev.* 8, 183–201. doi: 10.1007/s10567-005-6664-5

Williams, J., Murray, A., and Whelan, C. T. (2014). Multi-dimensional deprivation among 9-year-olds in Ireland: an analysis of the growing up in ireland survey. *Child Ind. Res.* 7, 279–300. doi: 10.1007/s12187-013-9215-5

Wish, N. B. (1986). Are we really measuring the quality of life? *Am. J. Econ. Soc.* 45, 93–99. doi: 10.1111/j.1536-7150.1986.tb01906.x

Woerner, W., Becker, A., and Rothenberger, A. (2004). Normative data and scale properties of the German parent SDQ. *Eur. Child Adoles. Psychiatry* 13 (Suppl. 2) II 3–10. doi: 10.1007/s00787-004-2002-6

Youth in Mind (2014). *SDQ: Information for Researchers and Professionals About the Strengths and Difficulties Questionnaire.* Available online at: http://www.sdqinfo.com/

APPENDIX A

Hand Clap Task

The Pencil Tapping task, which measures working memory, attention, and inhibitory control in 4 years or older children, from Luria-Nebraska Battery (Golden et al., 1979), was modified into hand clap task in 4.5-year DCW, in which children were requested to clap rather than tap a pencil as it is easier to follow by children. This task is sensitive to development, and has been executed in a diverse sample of children (Diamond and Taylor, 1996). The task comprised 16 trials in which children were meant to clap once if the assessor clapped twice and vice versa. The participants were assigned a score of 1 for a correct response while 0 was given for a wrong response. The scores were later aggregated. The above-mentioned abilities are implicated in learning behavior and social competence (Denham et al., 2012).

Pragmatic Language

Pragmatic language is an individual's ability to communicate over and above their vocabulary. It was assessed through Parent Rating of Oral Language and Literacy (PROLL; Dickinson et al., 2001), which is an adapted version of Teacher Rating of Oral Language and Literacy. This ability was measured through the following five questions at 4.5-year DCW.

- Which of the following best describes {Name}'s pattern of asking question?
- How often does {Name} try out new word?
- Which of the following best describes {Name}'s ability to communicate personal experiences in a clear and logical way?
- How often is {Name} understandable when speaking to adults other than you or other family?
- Which of the following best describes {Name}'s ability to communicate if {He/She} is not first understood?

Exploratory Factor Analysis with Direct Oblimin rotation showed this items loading onto one factor with loading ranging from 0.45 to 0.54 ($\alpha = 0.61$). Therefore, the responses on these items were averaged to give a single score for pragmatic language.

School Readiness

School readiness was calculated using seven relevant questions drawn from data collected at 6-year DCW from the mother/caregiver:

- I think that my child/children feel like their school is a good place to be.
- I think that my child/children feel like they belong in their school.
- I think that my child/children are happy in their school.
- I think that my child/children can mix with other children well at school.
- I think that my child/children have the reading and writing skills necessary at school.
- I think that my child/children find it easy to be left at school each morning.
- I think that my child/children are independent enough to cope with school.

Each question had seven response options (1 = *Strongly Disagree*" to "5 = *Strongly Agree*," with additional options of "*Don't know*" and "*Prefer not to say*," which were coded as missing). Exploratory Factor Analysis with Direct Oblimin rotation showed these questions loading onto one factor with loadings ranging from 0.72 to 0.48 ($\alpha = 0.80$). The responses on these questions were averaged to give a single score of readiness for school.