Volume 30, Number 2, 2022
ISSN 1061-3749

JNM

Journal of Nursing Measurement

An Examination of Threshold Setting in Social Emotional Measurement

Jianjun Wang, PhD

California State University, Bakersfield, 9001 Stockdale Highway, Bakersfield, CA 93311

Background and Purpose: Thresholds are used by nurses in social emotional screening. A purpose of this study is to examine fluctuation of the threshold pattern in a widely-used instrument, Ages and Stages Questionnaire-Social Emotional, Second Edition (ASQ:SE-2). **Methods:** Threshold settings are investigated against an irreversible pattern of child growth from research literature. Empirical studies are conducted on an issue of 6,039 missing cases in the ASQ:SE-2 normative sample that undermines data quality for the cutoff score configuration. **Results:** A Bayesian estimate has been suggested to improve the threshold at age 3 with an asymptotically unbiased cutoff score. Given the coverage of ASQ:SE-2 for children from less than 6 months to 72 months for mental health referrals, this study is particularly relevant to the practice of social emotional screening performed by pediatric nurses, school nurses, and nurse practitioners at mental health hospitals. **Conclusion:** To address the violation of Classical Test Theory and Item Response Theory in the ASQ:SE-2 threshold settings, consistent statistical imputations are needed to maintain monotonicity of the cutoff score patterns that are aligned with the irreversible trend of child growth, as well as the well-established national standards for missing data examination.

Keywords: social-emotional assessment; non-monotonic pattern; data imputation; threshold setting

Proper use of screening thresholds has been found to impact effectiveness of daily nursing practice (Graves et al., 2016). Social emotional screening for children has drawn more attention because 90% of human brain development occurs in the first five years. This field may attract interest of both *pediatric nurse* and *school nurse* as the federal government attempts to include preschool in the U.S. compulsory education (see Loiaconi, 2021). Since "Social and emotional development are as important to health and learning" (Nelson et al., 2013, p. 317), nurse practitioners are expected to support child screening and case referrals to mental health hospitals during the process of child growth.

To address the service need, "It is necessary to develop and document validated and appropriate methods of screening for children's early mental health problems" (Borg et al., 2014, p. 1). In the United States, the Ages & Stages Questionnaire: Social-Emotional (ASQ:SE) is one of a few recommended screening tools for monitoring children's social-emotional development during months 6–60 (American Academy of Pediatrics, 2001). In 2015, the ASQ:SE instrument was updated to its second edition (i.e., ASQ:SE-2) to expand the age coverage in the periods before month 6 and during 60–72 months. This change has allowed the use of ASQ:SE-2 to "assist with screening children during kindergarten and entry to school" (Squires et al., 2015a, p. 183).

http://dx.doi.org/10.1891/JNM-2021-0023

The critical role of threshold setting is reflected by its impact on mental health referrals. Accordingly, threshold comparisons have been emphasized in nurse practitioner preparation for mental health hospitals (Smith et al., 2004). Whenever the result of social emotional screening is above an age-specific threshold, children should be further assessed for potential referrals (Pethe et al., 2020). Given importance of the instrument application in child protection, a purpose of this research is to examine consistency of social emotional screening, per influence of the threshold identification, between ASQ:SE and ASQ:SE-2. The study is needed because ASQ:SE is still in use after the release of ASQ:SE-2. Consensus has yet to be established on the instrument replacement (Bush et al., 2018) and "some even argue that transitions should not occur until there is sufficient time to conduct research on newer measures" (Kamara et al., 2020, p. 362).

LITERATURE REVIEW

This investigation has a global impact because of the instrument use in many countries, including Brazil (Anunciação et al., 2019), China (Xie et al., 2019), Colombia (Bernal, 2019), Ethiopia (Abessa, 2019), Norway (Polte, 2019), Pakistan (Sikander et al., 2019), Scandinavia (Marks et al., 2020), Swenden (Vaezghasemi et al., 2019), Turkey (Zengin et al., 2019), and the United States (Marks et al., 2020). According to the ASQ:SE and ASQ:SE-2 technical reports (Squires et al., 2002; Squires et al., 2015a), cutoff scores of the two instruments are extracted in Table 1 across months 6, 12, 18, 24, 30, 36, 48, and 60. Built on these age-specific thresholds, "Whether children are referred for further evaluation and possible intervention or receive no further specialized services can have a lifelong impact for young children and their families" (Yovanoff & Squires, 2006, p. 48).

Inspection of Table 1 uncovers a difference in the cutoff score patterns surrounding month 36. While cutoff scores of ASQ:SE increase monotonically during child growth, the ASQ:SE-2 cutoff scores fluctuate with a peak at month 36—After age 3, the cutoff score drops from 105 to a smaller number in 48th and 60th months.

The anomaly of cutoff score variation has made the ASQ:SE-2 result much less sensitive to developmental issues at 36th month. As a result, a screening score of 91 could lead to mental health referral because it was larger than the cutoff score of 85 in months 30 and 48 (Table 1). However, 91 is below the threshold of 105 at 36th month. Even with no change in the screening score, the referral decision could be reversed due to the threshold fluctuation. This example illustrates less sensitivity of ASQ:SE-2 with false negatives for detecting mental health issues around age 3.

The issue is not trivial because "The alteration of a screening cutoff score by one or two raw score points might significantly impact a referral decision for a child and family and might determine very different developmental trajectories for the child" (Yovanoff & Squires, 2006, p. 48). A 20-point drop is demonstrated in the cutoff scores between 105 at 36th month and 85 in the neighboring 30th and 48th months. Since the gap is larger than the threshold difference between any other consecutive age groups in Table 1, more concerns could be raised on reliability of the assessment outcomes at age 3.

Selection of screening instrument, particularly for young children, has been an active front of nursing research (Borg et al., 2014). More recently, many countries attached great importance to social emotional assessment at 36th month. For instance, Vaezghasemi et al. (2019) reported that 40 Child Health Care (CHC) centers in northern Sweden had about 3000 children at age 3 each year. Despite availability of ASQ:SE-2, their instrument choice

of ASQ:SE has avoided the issue of non-monotonic threshold pattern that might cause inconsistent referral decisions.

To track down the issue, cutoff scores are empirically derived from a normative sample that is designed to represent the target population. Thus, inaccuracy of the cutoff scores may hinge on issues of the normative sample composition (Yovanoff & Squires, 2006). The normative sample for ASQ:SE-2 contains 16,424 children (Squires et al., 2015a), much larger than the original ASQ:SE sample of 2,861 children with missing data in 382 cases (Squires et al., 2002). Moreover, no information was gathered on developmental status of 6,039 children in ASQ:SE-2 due to missing responses from parents/primary caregivers (Squires et al., 2015a, p. 190). Besides a substantial increase of the sample size, the proportion of missing data also increased considerably in ASQ:SE-2.

Per data quality standards of National Center for Education Statistics (NCES) (2003), "When unit nonresponse is high, nonresponse bias analysis must be conducted at the unit level to determine whether or not the data are missing at random and to assess the potential magnitude of unit nonresponse bias" (p. 88). Hence, the literature review indicates strong needs for investigating the missing data impact on cutoff score settings that were ignored in the ASQ:SE-2 technical report (Squires et al., 2015a).

Based on the existing literature, this study further adopts an exploratory approach to seek solutions to the missing data issue behind threshold configurations. To streamline the result reporting, the remainder of this article is organized as follows. First, queries are made on the cutoff score pattern under well-established Classical Test Theory (CTT) and Item Response Theory (IRT). The missing data impacts are subsequently assessed using a *context, input, process* and *product* (CIPP) model. Based on the cutoff score patterns between ASQ:SE and ASQ:SE-2, more investigation is pursued at the item level to compare instrument structures across different age groups. As a result, a Bayesian estimate is adduced from data imputation to improve the threshold at age 3 with a good control of asymptotical bias. Implication of the research findings is addressed at end of the article to clarify meaningfulness of this study to nursing practitioners.

CTT AND IRT EVALUATIONS OF THE CUTOFF SCORE PATTERN

Anunciacao, Squires et al. (2019) testified, "In psychometrics, both CTT and IRT have specific methods, concepts, and statistical techniques to address the variety of a test and its measurement problems. Nowadays, both frameworks are seen as complementary" (p. 5). In CTT, it is postulated that the observed score (x) equals a true score (T) plus an error item (E) (see Novick, 1966). Although the exact T value is unknown, individual scores are expected to surround the true score if measurement errors occur randomly. Since children grow continuously on the time dimension, the ASQ:SE-2 cutoff score is expected to change along with the maturation process. Therefore, the trend should be monotonic to match the irreversible child development.

Perversely, the cutoff score fluctuation, as indicated by the ASQ:SE-2 technical report (Squires et al., 2015a), seems to imply that children may grow and shrink interchangeably at an age near 36th month. However, the fluctuation does not match the true score growth during child development under the CTT paradigm.

In addition, IRT (a.k.a. modern mental test theory) models the relationship between performance outcomes and test taker abilities (Lord, 1980). As Yang and Kao (2014) cautioned,

TABLE 1. Cutoff Scores of ASQ:SE and ASQ:SE-2 for Mental Health Referral

Month	ASQ:SE	ASQ:SE-2
2		35
6	45	45
12	48	50
18	50	65
24	50	65
30	57	85
36	59	105
48	70	85
60	70	95[a]

Sources: https://agesandstages.com/wp-content/uploads/2015/09/ASQSE-2-Technical-Appendix_web.pdf; http://archive.brookespublishing.com/documents/asq-se-technical-report.pdf
[a]ASQ:SE-2 extends threshold 95 for children in months 60–72.

> One assumption of IRT is monotonicity, which is best displayed on a graph as a curve shaped like an "S" between the latent trait level on the *X*-axis and the probability of a more extreme response on the item (e.g., a question about depression) on the *Y*-axis. (p. 172)

As shown in Table 1, the threshold drop after age 3 has reversed the initial trend of threshold increase. Hence, the ASQ:SE-2 assessment outcomes cannot be supported by IRT.

In summary, "Both IRT and classical test theory (CTT) are concerned with the capacity of an instrument to measure the intended construct and both can be useful for quantitatively assessing items and scales" (Anunciacao et al., 2019, p. 3). The true score configuration in CTT is expected to match the trajectory of child growth that stipulates no fluctuation of cutoff scores surrounding 36th month. For older children, a drop of the cutoff score after age 3 also violates the monotonicity assumption of IRT. Because the threshold configuration was based on the normative sample, data representativeness is subject to further investigation given the missing value acknowledgment in the ASQ:SE-2 technical report (Squires et al., 2015a).

A MODEL-BASED APPROACH FOR THE CUTOFF SCORE EXAMINATION

In the 21st century, "Practice nurses have become increasingly involved in providing mental health services" (Robinson, 2003, p. 13). To address quality assurance of social emotional screening for mental health referrals, a comprehensive model is adopted to articulate phases of *context, input, process,* and *product* (CIPP) in a coherent framework for instrument evaluation. With the cutoff score as a *product* for the ASQ:SE-2 score comparison, the impact of missing data needs to be tracked in the *process* of threshold configuration. In addition, screening items represent the *input* ingredients that compose "the dimensions of self-regulation, compliance, communication, adaptive functioning, autonomy, affect, and interaction with people" in social emotional assessment (Folger et al., 2019, p. 4). Hence, further analyses should be directed to item differences in the instrument composition.

Investigation of the sample representation is also inseparable from an examination of the demographic data through comparisons with the corresponding information in the broad population *context*. Stufflebeam and Coryn (2014) delineated CIPP features for model-based investigation.

Process of Cutoff Score Setting

From the *process* point of view, *products* of social emotional assessment depends on the procedure of cutoff score computing through a receiver operating characteristic (ROC) analysis to optimize the threshold choice with a balanced consideration of several important measures, including sensitivity and specificity indices (Malmberg et al., 2003). In the ASQ:SE-2 technical report (Squires et al., 2015a), an ROC-based result table is titled "Range of points, medians, interquartile ranges, receiver operating characteristic (ROC) cutoffs, and percentages identified by ASQ:SE-2 interval (N = 16,424)" (p. 195), which seems to suggest data availability from 16,424 children in the ROC computing.

Unfortunately, the cutoff score configuration was not grounded on the entire normative sample because of missing data in 6,039 cases. The missing data issue is revealed from reviewing another footnote on a different page of the ASQ:SE-2 technical report, i.e., "Developmental status data are missing for 6,039 children" (Squires et al., 2015a, p. 191). Given the dependency of cutoff score configuration on developmental status, Squires et al. (2015a) acknowledged the "Number of questionnaires by developmental status for the ASQ:SE-2 normative sample (N = 10,385)" (p. 191). Thus, the achieved sample size is 10,385. Altogether the numeric comparison reconfirms the missing of 6,039 cases in the cutoff score configuration, i.e., 16,424—10,385 = 6,039. Apparently, the proportion of missing data has surpassed the negligible level set by the NCES (2003) data standards.

Differences in the Item Composition

In comparison, the missing case count was 382 in ASQ:SE (Squires et al., 2002), much less than 6,039 in ASQ:SE-2. When the magnitude of missing information was not high, no violation of the monotonicity pattern has been observed in ASQ:SE cutoff scores (see Table 1). Since ASQ:SE-2 was developed from ASQ:SE, item compositions can be tracked between the two instruments to examine the impact of fluctuation pattern in cutoff scores. Without a thorough examination, nurses may overlook "the subtle cues of childhood stress that prevent them from intervening in a timely and appropriate manner for optimal effectiveness" (Gerdner et al., 2005, p. 222).

The ASQ:SE items are available from Squires et al. (2002) and ASQ:SE-2 items are released by Squires et al. (2015b). For mental health referrals at 36th month, the ASE:SE instrument includes 34 items, and all of them have been inherited by the ASQ:SE-2 instrument (Table 2). In addition, ASQ:SE-2 incorporates four new items to increase the total number of items to 38. These items are:

- *Does your child try to show you things by pointing at them and looking back at you?*
- *Does your child pretend objects are something else? For example, does he pretend a banana is a phone?*
- *Does your child wake three or more times during the night?*
- *Is your child too worried or fearful? If "sometimes" or "often or always," please*

Based on the instrument difference, one may question whether the four new items caused a jump of the cutoff scores from 59 in ASQ:SE to 105 in ASQ:SE-2 at 36th month

TABLE 2. Overlap of Screening Items for ASQ:SE (month 36) and ASQ:SE-2 (months 30, 36 & 48)

Item sequence number				
ASQ:SE-2 (36th month)	**ASQ:SE (36th month)**	**ASQ:SE-2 (48th month)**	**ASQ:SE-2 (30th month)**	**Label**
1	**1**	1	1	Look when talk
2	**2**	5	2	Like hug/cuddle
3	**3**	3		Play with adult
4	**4**	2	3	Cling to you
5	**5**	4	14	Calm down
6	**6**	6	7	Too friendly
7	**7**	7	8	Settle down
8	**8**	20	22	Move on
9	**9**	14	5	Seem happy
10	**10**	9	13	Interested in things
11	**11**	13	12	Do what is asked
12	**12**	16		More active
13	**13**	18	11	Stay with activities
14	**14**	12	16	Enjoy mealtimes
15	**15**	11	15	Eating problem
16	**16**	15	18	Enough sleep
17	**17**	17		Use words to tell
18	**18**	24	20	Follow directions
19	**19**	8	9	Cry for long time
20	**20**	21	21	Check you are near
21	**21**	22	10	Upset for being stopped
22	**22**	23	25	Hurt self on purpose
23	**23**	26	23	Stay away from danger
24	**24**	25	24	Damage things on purpose
25	**25**	19		Describe feelings
26	**26**	27		Name a friend
27	**27**	29		Others like to play with him/her
28	**28**	30	26	He/she like to play with others
29	**29**	31	27	Try to hurt others
30	**30**	32		Unusual interest in sexual stuff
31			**28**	Point things to show you
32			**30**	Pretend objects as other things
33		33	**31**	Wake 3 or 4 time at night
34		34	**32**	Too fearful
35	**31**	36	33	Concerns shared by others
36	**32**		34	Eat, sleep, toilet concern
37	**33**		35	Does anything worry you
38	**34**		36	What do you enjoy about child

Note. Bold font is used to show augmentation of four items in instrument upgrading from ASQ:SE to ASQ:SE-2 in the 36th month screening.

(see Table 1). In other words, if 59 is an appropriate cutoff score from the original ASQ:SE instrument that contains 34 items, can the four new items profoundly cause a 46-point (i.e., 105–59 = 46) increase of the cutoff score in ASQ:SE-2?

As shown in Table 2, these items not only existed as *items 31–34* for ASQ:SE-2 assessment at age 3, but also appeared as *items 28, 30, 31, and 32* in the 30th month instrument. In addition, they also emerged as *items 26, 27, 28, and 30* in the ASQ:SE-2 instruments for 24th month (Squires et al., 2015b), and two of the items were labeled as *items 33 and 34* in the ASQ:SE-2 assessment for 48th month (Table 2). Hence, augmentation of the four new items in ASQ:SE-2 is unlikely to cause the score fluctuation exclusively at 36th month. Based on the item existence in other months, the corresponding cutoff score increase could have wiped out the non-monotonicity issue on the time dimension. Therefore, the cutoff score fluctuation is unlikely to result from item changes. However, it is possible that skewed missing cases in the ASQ:SE-2 normative sample may undermine the data representation for the age-specific child population.

Skewness in Population Representation

Mental health referral is a major public health service in nurse preparation (Pryjmachuk et al., 2011). In developing ASQ:SE-2, Squires et al. (2015a) maintained that "An attempt was made to stratify the normative sample so that children/families would be representative of the U.S. population in terms of race/ethnicity, geographic region, parent education and income, and gender of children" (p. 186). However, they also admitted that "our demographic data were collected in different categories" (Squires et al., 2015a, p. 189).

For these variables using the same categorizations from the Census Bureau, results show skewed representation of the sample data for the population. In particular, the ASQ:SE-2 normative sample has 9,460 males and 6,934 females (Squires et al., 2015a). The oversampling of males does not provide a balanced representation for the target population that has more females than males from the decennial census (U.S. Census Bureau, 2010). In terms of parent education, Figure 1 shows 3.7% of the normative sample below high school diploma. Meanwhile, 56.7% of the normative sample comes from families with parents achieving a *4-year college degree or above* (Figure 1). Because of the linkage

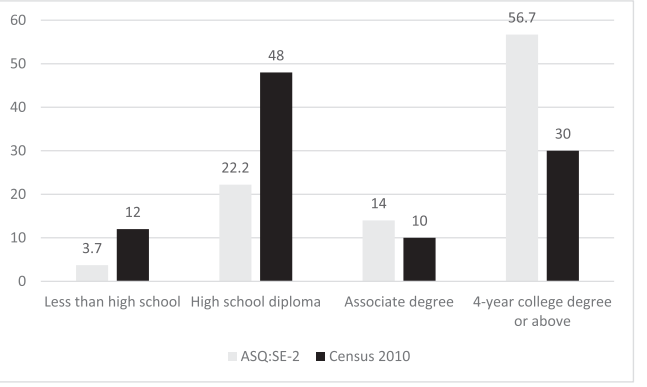

Figure 1. Percent of parent highest education across different levels.

Source: ASQ:SE-2 Technical Report (Squires et al., 2015a).

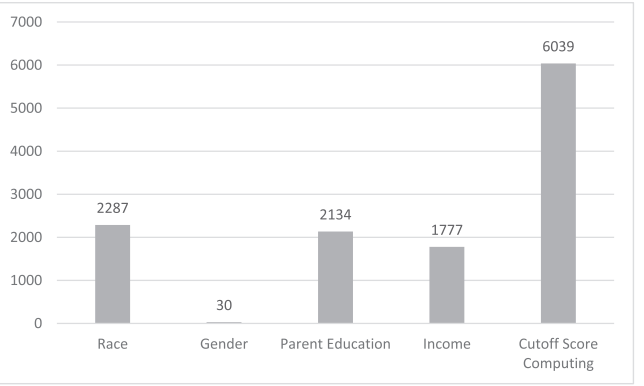

Figure 2. Missing data in the normative sample for ASQ:SE-2 computing.

between education and income, parents are also skewed toward a category of earning *more than $40,000* (Squires et al., 2015a, p. 190).

In this context, Figure 2 shows the total count of 6,228 missing cases across the *race, gender, parent education,* and *income* dimensions. The largest count is in the *race* category that contains 2,287 missing cases. In comparison, it is still much less than the count of 6,039 cases with missing information on developmental status. The skewed demographic data may undermine representativeness of the normative sample in ASQ:SE-2.

DISCUSSION

In reporting the ASQ:SE-2 results, Squires et al. (2015a) acknowledged, "The general trend of increasingly higher scores as children develop is reflected in median scores, except at 48 and 60 months" (p. 194). To date, the ASQ:SE-2 researchers did not fix the non-monotonicity trajectory of child growth during ages 0–5, nor did they attempt to impute missing data for improvement of the cutoff score pattern.

In coping with missing data, Shlomo (2015) offered potential solutions from *imputation* and *post-stratification* perspectives. Although it is unrealistic to conduct multiple imputations of the missing data that exceed one third of the normative sample size (Wang & Johnson, 2019), the nearest neighbor imputation is another feasible method. Rancourt et al. (1994) reported that its estimates were asymptotically unbiased, which fit the large sample context of ASQ:SE-2. Adjacent to 36th month is a cutoff score of 85 for the neighboring months 30 and 48 (Table 1). The nearest neighbor imputation consistently supports replacement of 105 with 85 as the cutoff score at 36th month to regain the desired pattern of monotonicity.

On the other hand, the post-stratification process is built on a premise that a naive or raw estimate can be improved by combining it with other information to strengthen the population representation. Incorporation of the extra conditions or priors is also considered a Bayesian approach. In particular, "The informative prior is reasonable to use if one has real prior information from a previous similar study" (Patetta, 2017, Ch. 1, p. 12). For ASQ:SE-2, its predecessor is ASQ:SE. Hence, the monotonicity feature of ASQ:SE can be incorporated to adjust cutoff scores in ASQ:SE-2. With a common cutoff value of 85 in the neighboring months 30 and 48 (Table 1), 85 is the only Bayesian estimate to avoid

the trajectory fluctuation, and thus, maintain monotonicity of the cutoff scores surrounding 36th month.

In conclusion, the need for correcting non-monotonicity is not only based on CTT and IRT models of social emotional assessment, but also aligned with the irreversible growth trajectory in early childhood development (Piaget, 1952). The incorporation of external information, such as ASQ:SE and data imputation, to improve the result from a controversial sample has been termed as *shrinkage* estimation (Zhao et al., 2010). Although shrinkage estimators are more accurate (Shor et al., 2007), Shadish et al. (2013) cautioned potential generation of competing values. However, it is not an issue here because adjustments from both *imputation* and *post-stratification* approaches support a cutoff score of 85 as the single threshold choice to improve the ASQ:SE-2 screening results at 36th month.

The threshold recommendation is designed to ensure alignment of the social emotional screening outcomes with the irreversible growth trajectory of early childhood development. This investigation is important because "Social and emotional difficulties in the first 5 years of life can have long-term consequences in that they affect learning, school achievement, and peer relationships in later childhood" (Breitenstein et al., 2007, p. 313).

As a general strategy, this research features both exploratory and confirmatory investigations. While the exploratory part disentangles the fluctuation of ASQ:SE-2 cutoff scores that has never been studied before, the model-based approach, including uses of CTT, IRT, and CIPP paradigms, conforms to the professional practice for result triangulation. Given the broad market for ASQ:SE-2 adoption after COVID-19, the suggestion for threshold improvement represents an important progress to safeguard accurate mental health referrals. Otherwise, numerous children could have been misjudged by the problematic thresholds from social emotional screenings according to the well-disseminated ASQ:SE-2 technical report (Squires et al., 2015a).

IMPLICATION FOR NURSING PRACTICE

In the 21st century, social emotional screening has become a primary task for nurse practitioners because "More than 90 per cent of patients with mental health problems are treated in primary care" (Robinson, 2003, p. 13). Therefore, significance of this study can be clarified through elaboration of the inquiry impact for different nursing practitioners, including pediatric nurses, school nurses, mental health nurses, and researchers from the Doctor of Nursing Practice (DNP) programs in need of understanding measurement consistency in social emotional screening. The key stakeholder identification naturally hinges on the feature of ASQ:SE-2 that extends the threshold setting of ASQ:SE from month 60 to month 72 (Squires et al., 2015a). Accordingly, an accurate interpretation of the overall screening patterns is not only relevant to pediatric nurses for infant and/or toddler care, but also germane to school nurses in an elementary school setting. In addition, nurses at mental health hospitals need to understand the threshold setting as a foundation for case referrals to treat social emotional issues.

In the existing U.S. education system, children in 60–72 months are kindergartners experiencing an unprecedented transition from home to school. The change is inevitably accompanied by additional stress to build new relationships with unfamiliar peers in a classroom setting (Commissioner for Children and Young People, 2011). To smooth the adjustment, it is important for the school nurse to detect mental health problems related to social and emotional stressors (Shannon et al., 2010). Therefore, well-rounded school

readiness is not only confined within the cognitive domain, but also extends to mental health wellbeing. The transition challenge is universal, and lacking valid and reliable instruments for identifying social and emotional delays in young children is a worldwide issue (Xie et al., 2019).

Prior to school entry, accurate social emotional screening is also important for nurses in primary care. In general, "Primary care services were provided to the children by pediatric nurse practitioners, pediatricians, registered nurses and licensed practical nurses" (Hallas et al., 2017, p. 35). For pediatric nurse practitioners serving maltreated infants and toddlers, addressing social emotional problems is integral to the childcare commitment (Hash, 2018). In comparison to preschoolers at age 3, kindergartners at age 5 or above have a lower cutoff score in ASQ:SE-2 (see Table 1). Consequently, mis-referrals might occur to send relatively more kindergartners to mental health hospitals due to the threshold fluctuation revealed in this study.

Social-emotional development and research related to the interventions are of the interest of mental health clinicians, family therapists, nurses, and parent educators (Erturk, 2019). In particular, DNP researchers need to engage in family interventions to "determine the effect on maternal confidence and toddler social-emotional development" (Hallas et al., 2017, p. 39). Senyuva et al. (2019) stressed importance of assessment skill development in nursing education to facilitate screening result interpretation. While clarification of the result implication has substantiated relevance of this study in nursing research, practice, and education, more investigations can be conducted in the future to strengthen child wellbeing from the perspectives other key stakeholders, such as mental health clinicians, family therapists, and parent educators.

REFERENCES

Abessa, T., Worku, B., Wondafrash, M., Girma, T., Valy, J., Lemmens, J., Bruckers, L., Kolsteren, P., & Granitzer, M. (2019). Effect of play-based family-centered psychomotor/psychosocial stimulation on the development of severely acutely malnourished children under six in a low-income setting: A randomized controlled trial. *BMC Pediatrics*, *19*(1), 336–336.

American Academy of Pediatrics. (2001). Developmental surveillance and screening of infants and young children. *Pediatrics*, *108*, 192–196.

Anunciacao, L., Squires, J., Clifford, J., & Landeira-Fernandez, J. (2019). Confirmatory analysis and normative tables for the Brazilian Ages and Stages Questionnaires: Social-Emotional. *Child Care Health Development*, *45*(3), 387–393. https://doi.org/10.1111/cch.12649

Bernal, R., Attanasio, O., Peña, X., & Vera-Hernández, M. (2019). The effects of the transition from home-based childcare to childcare centers on children's health and development in Colombia. *Early Childhood Research Quarterly*, *47*, 418–431.

Borg, A., Salmelin, R., Joukamaa, M., & Tamminen, T. (2014). Cutting a long story short? The clinical relevance of askingparents, nurses, and young children themselves to identify children's mental health problems by one or two questions. *The Scientific World Journal*, *2014*, 286939. https://doi.org/10.1155/2014/28693931

Breitenstein, S. M., Gross, D., Ordaz, I., Julion, W., Garvey, C., & Ridge, A. (2007). Promoting mental health in early childhood programs serving families from low-income neighborhoods. *Journal of the American Psychiatric Nurses Association*, *13*(5), 313+. https://link.gale.com/apps/doc/A170595126/AONE?u=googlescholar&sid=googleScholar&xid=a57e4219

Bush, S. S., Sweet, J. J., Bianchini, K. J., Johnson-Greene, D., Dean, P. M., & Schoenberg, M. R. (2018). Deciding to adopt revised and new psychological and neuropsychological tests: An interorganizational position paper. *The Clinical Neuropsychologist*, *32*, 319–325. https://doi.org/10.1080/13854046.2017.1422277

Commissioner for Children and Young People (2011). *Report of the Inquiry into the Mental Health and Wellbeing of Children and Young People in Western Australia* (Parliamentary). Western Australia: Author. Available at: http://www.ccyp.wa.gov.au/

Erturk, B. (2019). *Experimental examination of the effects of parent-implemented select intervention on social-emotional development of infants and toddlers* (Order No. 27668400). Available from ProQuest Dissertations & Theses Global: The Humanities and Social Sciences Collection. (2383535605).

Folger, A., Ding, L., Ji, H., Yolton, K., Ammerman, R., Van Ginkel, J., & Bowers, K. (2019). *Neonatal NR3C1 methylation and social-emotional development at 6 and 18 months of age*. https://doi.org/10.3389/fnbeh.2019.00014.

Graves, J., Hodge, C., & Jacob, E. (2016). Depression, anxiety, and quality of life in children and adolescents with sickle cell disease. *Pediatric Nursing*, *42*(3), 113–144.

Gerdner, L., Buckwalter, K., & Hall, G. (2005). Temporal patterning of agitation and stressors associated with agitation: Case profiles to illustrate the progressively lowered stress threshold model. *Journal of American Psychiatric Nurses Association*, *11*(4), 215–222. https://doi.org/10.1177/1078390305281178

Hallas, D., Koslap-Petraco, M., & Fletcher, J. (2017). Social-emotional development of toddlers: Randomized controlled trial of an office-based intervention. *Journal of Pediatric Nursing*, *33*, 33–40. https://doi.org/10.1016/j.pedn.2016.11.004

Hash, J. B. (2018). *Adverse childhood experiences, sleep, and social emotional development among infants and toddlers from families involved with child protective services* (Order No. 10748529). Available from ProQuest Dissertations & Theses Global: The Humanities and Social Sciences Collection. (2034397105).

Kamara, D., Walton, K., & Witwer, A. (2020). Socioemotional and autism spectrum disorder screening for toddlers in early intervention: Agreement among measures. *Journal of Early Intervention*, *42*(4), 359–380. https://doi.org/10.1177/1053815119880607

Loiaconi, S. (2021). *Biden eyes universal pre-K plan, but potential challenges loom*. https://fox11online.com/news/nation-world/biden-eyes-universal-pre-k-plan-but-potential-challenges-loom

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Lawrence Erlbaum Associates.

Malmberg, M., Rydell, A. M., & Smedie, H. (2003). Validity of the Swedish version of the strengths and difficulties questionnaire. *Nordic Journal of Psychiatry*, *57*, 357–363.

Marks, K., Madsen, N., & Wilson, P. (2020). Comparative use of the Ages and Stages Questionnaires in the USA and Scandinavia: A systematic review. *Developmental Medicine & Child Neurology*, *61*, 419–430.

National Center for Education Statistics (NCES). (2003). *NCES statistical standards*. https://nces.ed.gov/pubs2003/2003601.pdf.

Nelson, H., Kendall, G., & Shields, L. (2013). Children's social/emotional characteristics at entry to school: Implications for school nurses. *Journal of Child Health Care*, *17*(3), 317–331. https://doi.org/10.1177/1367493512461458

Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, *3*(1), 1–18.

Patetta, M. (2017). *Bayesian analyses using SAS*. SAS Institute.

Pethe, K., Maldonado-Soto, A., Saxena, J., Blanck, E., Lingras, K., & Aratani, Y. (2020). The relationship between linkages to behavioral health services in pediatric primary care and reductions in non-urgent emergency department visits among vulnerable children. *The Journal of Behavioral Health Services & Research*, *47*(3), 377–387.

Piaget, J. (1952). *The origins of intelligence in children*. International University Press.

Polte, C., von Junge, C., Soest, T., Seidler, A., Eberhard-Gran, M., & Garthus-Niegel, S. (2019). Impact of maternal perinatal anxiety on social-emotional development of 2-year-olds, a prospective study of Norwegian mothers and their offspring. *Maternal and Child Health Journal*, *23*(3), 386–396.

Pryjmachuk, S., Graham, T., Haddad, M., & Tylee, A. (2011). School nurses' perspectives on managing mental health problems in children and young people. *Journal of Clinical Nursing*, *21*, 850–859. https://doi.org/10.1111/j.1365-2702.2011.03838.x

Rancourt, E., Särndal, C., & Lee, H. (1994). Estimation of the variance in the presence of nearest neighbor imputation. *ASA Proceeding Section on Survey Research Methods*. 888–893.

Robinson, F. (2003). Mental health issues in primary care. *Practice Nurse*, *25(3), 13*, 15.

Shadish, W., Rindskopf, D., Hedges, L., & Sullivan, K. (2013). Bayesian estimates of autocorrelations in single-case designs. *Behavior Research Methods*, *45*, 813–821.

Shannon, R., Bergren, M., & Matthews, A. (2010). Frequent visitors: Somatization in school-age children and implications for school nurses. *The Journal of School Nursing*, *26*(3), 169–182. https://doi.org/10.1177/1059840509356777

Shlomo, N. (2015). Handling missing data, statistical data editing and imputation. https://ec.europa.eu/eurostat/cros/system/files/SpringSchool_SDE_2015.pdf

Shor, B., Bafumi, J., Keele, L., & Park, D. (2007). A Bayesian multilevel modeling approach to time-series cross-sectional data. *Political Analysis*, *15*(2), 165–181.

Sikander, S., Ahmad, I., Bates, L., Gallis, J., Hagaman, A., O'Donnell, K., Turner, E., Zaidi, A., Rahman, A., & Maselko, J. (2019). Cohort Profile: Perinatal depression and child socioemotional development; the Bachpan cohort study from rural Pakistan. *BMJ Open*, *9*(5), e025644–e025644. https://doi.org/10.1136/bmjopen-2018-02564

Smith, M., Gerdner, L., Hall, G., & Buckwalter, K. (2004). History, development, and future of the progressively lowered stress threshold: A conceptual model for dementia care. *Journal of American Geriatric Society*, *52*(10), 1755–1760. 10.1111/j.1532-5415.2004.52473.x. PMID: 15450057.

Squires, J., Bricker, D., & Twombly, E. (2002). *Technical Report on ASQ: SE*. http://archive.brookespublishing.com/documents/asq-se-technical-report.pdf

Squires, J., Bricker, D., Twombly, E., Murphy, K., & Hoselton, R. (2015a). *ASQ: SE-2 technical report*. Brookes Publishing.

Squires, J., Bricker, D., & Twombly, E. (2015b). *Ages & Stages Questionnaires. Social-Emotional (Second Edition)*. Brookes Publishing.

Stufflebeam, D., & Coryn, C. (2014). *Evaluation theory, models, and applications*. Jossey-Bass.

Senyuva, E., Kucuk Yuceyurt, N., & Ulupinar, S. (2019). Does participation of nursing students in social activities affect their social emotional learning skills? *Nurse Education Today*, *76*, 78–84. https://doi.org/10.1016/j.nedt.2019.01.031

U.S. Census Bureau. (2010). *Age and sex composition: 2010*. https://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf

Vaezghasemi, M., Eurenius, E., Richter Sundberg, L., Ivarsson, A., Silfverdal, S., & Lindkvist, M. (2019). Social-emotional problems among Swedish three-year-olds: An item response theory analysis. *European Journal of Public Health*, *29*(Suppl 4), 367.

Wang, J., & Johnson, D. (2019). An examination of discrepancies in multiple imputation procedures between SAS and SPSS. *The American Statistician*, *73*(1), 80–88.

Xie, H., Bian, X., Chen, C., Squires, J., & Lu, P. (2019). Examining the convergent evidence of a parent-completed, Social-Emotional screening tool in China. *Journal of Child and Family Studies*, *28*(6), 1471–1480.

Yang, F., & Kao, S. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry*, *26*(3), 171–177. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4118016/pdf/sap-26-03-171.pdf

Yovanoff, P., & Squires, J. (2006). Determining cutoff scores on a developmental screening measure: Use of receiver operating characteristics and item response theory. *Journal of Early Intervention*, *29*, 48–62.

Zengin, P., Gharibzadeh, M., Ilter, E., Ozmert, E., Eryilmaz, S., Ozdemir, G., Karahan, S., Yalcin, E., Dogru, D., Kiper, N., & Ozcelik, U. (2019). Developmental and behavioral problems in preschool-aged primary ciliary dyskinesia patients. *European Journal of Pediatrics*, *178*(7), 995–1003.

Zhao, Y., Lee, A., & Barnes, T. (2010). On application of the empirical Bayes shrinkage in epidemiological settings. *International Journal of Environmental Research and Public Health*, *7*(2), 380–394.

Disclosure. The author has no relevant financial interest or affiliations with any commercial entities related to the subjects discussed within this article.

Acknowledgments. The author wish to thank First 5 Kern for revealing needs for this investigation. Also thank Dr. James Rodriguez for stressing importance of early childhood research.

Funding. The author(s) received no specific grant or financial support for the research, authorship, and/or publication of this article.

Correspondence regarding this article should be directed to Jianjun Wang, PhD, California State University, Bakersfield, 9001 Stockdale Highway, Bakersfield, CA 93311. E-mail: jwang@csub.edu