Check for updates

Development of Parent Guidelines for Parent-Performed Developmental Screening Tests

Sung Sil Rah[1], Soon-Beom Hong[2], and Ju Young Yoon[3]

[1]Institute of Behavioral Science in Medicine, Yonsei University College of Medicine, Seoul, Korea
[2]Division of Child and Adolescent Psychiatry, Department of Psychiatry, College of Medicine, Seoul National University, Seoul, Korea
[3]Research Institute of Nursing Science, Seoul National University, Seoul, Korea

Objectives: Most developmental screening tests have been built as parent-performed questionnaires. However, they often do not guide parents on how to answer the questionnaire. This study aimed to develop easily applicable parent guidelines.
Methods: We implemented the Delphi procedure with 20 panelists. The development of the initial questionnaire was based on the results of two surveys of parents and experts provided by a policy research report that investigated the item adequacy of the Korean Developmental Screening Test. Round one included 33 items comprising all possible measurements in six categories that were identified as difficult to understand or confusing. Round two merged and modified some items and included 32 items. We defined consensus as a median agreement value of one or less and convergence and stability values of 0.5 or less. The subjective usefulness of the parent guidelines was examined based on their previous test experiences.
Results: Consensus was reached after the second round, reflecting the items with the highest level of accuracy in each category. Of the 167 parents who participated in the survey, 113 (67.7%) affirmed the usefulness of the guidelines, while 10 (6.0%) answered that they were not useful. Items that recommended a different scoring strategy in answering the questionnaire from their previous measurements were found to be more useful by the parents.
Conclusion: The parent guidelines, composed of five bullet points, drew on the consensus of the experts. Further studies are required to assess whether these guidelines improve the accuracy of screening tests in clinical settings.

Keywords: Child; Delphi technique; Developmental disabilities; Diagnostic screening programs; Practice guidelines as topic.

INTRODUCTION

With the increasing prevalence of developmental disabilities (DDs) among young children worldwide [1-3], early identification of DDs has become even more critical for allowing the developmental trajectory of children with DDs with appropriate interventions [4-6]. Due to the high cost and insufficient number of specialists [2,7], however, access to developmental confirmatory tests is limited, which creates a delay in the diagnosis of DDs and causes missed opportunities [8]. As a result, developmental screening tests have emerged as part of an accurate and cost-effective health management system for young children at the national level, with an expectation of better prognosis for children at risk of DDs [9]. Many screening tests, such as the Denver Developmental Screening Test, the Ages and Stages Questionnaires (ASQ), and the Parent's

Evaluation of Developmental Status (PEDS), have been developed and implemented [10-12], and parent-administered screening tests are more popularly used according to the recommendation of the American Academy of Pediatrics [13]. One of the most widely used developmental screening tests is the ASQ, which is designed for periodic application with infants and children under six years old [12]. Since its development in the 1970s in the US [14], it has been translated into dozens of languages and used globally both in clinical and experimental settings [15,16]. Despite the broad use of developmental screening tests like the ASQ, however, some researchers question the effectiveness of these screening tests, as they have shown only moderate accuracy in previous studies. Sheldrick et al. [17] compared the accuracy of three parent-performed developmental screening tests—ASQ, PEDS, and the Survey of Well-being of Young Children (SWYC)—among a total of 642 children, ages 0 to 66 months. Unlike the estimates of specificity higher than 70% for all three screening tests, they presented low to moderate accuracy in sensitivity, from

23.5% for the ASQ in the younger group to 61.8% for the PEDS in the older [17]. Results of another study that analyzed nationwide population-based data revealed moderate levels of sensitivity of the Korean version of ASQ and the Korean Developmental Screening Test for Infants and Children (K-DST)—64.1% and 44.4%, respectively, for the lowest [18].

One suspected cause is the reliability of parents as test administrators. In a study by Sheldrick et al. [17], parents of 298 (20.5%) children found them to be positive on the ASQ, while 422 (29.0%) and 127 (8.8%) children were classified as positive on the PEDS and the SWYC, respectively; the concordance rate of two screening tests among the three ranged from 35% to 60%. In another study that examined agreement of the test results between the ASQ and the PEDS, 33% (20/60) of the test results disagreed; for those who received positive results from one or both of the tests, 69% (20/29) showed disagreement in the test results [19].

According to survey results in a policy research report that investigated difficulties in administering the K-DST, 40.2% (125/311) of the parents had difficulty answering the questionnaires due to failure in observing their child's performance (86, 27.7%) and the confusing or difficult questions (36, 11.6%) [20]. In addition, 35 of 85 (43.2%) professionals reported that parents could not understand the meaning of the questions and needed additional explanations [20]. Despite these reported difficulties, most of these screening tests do not give parents sufficient administration guidance [21-23]. The K-DST user's guide devotes only half a page to parental administration, and that only includes instructions to answer the question after they let their child perform the task, if they are not sure of the child's capability, and to mark as "can do it" if the child shows sufficient abilities even without witnessing the actual performance [21]. Other screening tests are not very different—the ASQ provides a 4-page flyer for parents that includes general instructions about the screening process along with the purpose and the expected benefits [23]; the PEDS provides guidelines not for parents but only for professionals to help score the results [24]. None of these guidelines provides scoring criteria on the scale, which may compromise diagnostic accuracy and cause a delay in identification of children with DDs, especially for those with mild developmental delay [17].

Therefore, this study aimed to 1) develop easily understandable guidelines that can help parents accurately administer a parent-performed developmental screening test and 2) evaluate the subjective usefulness of the developed guidelines.

METHODS

We developed the guidelines based on the K-DST. The K-DST is a parent-performed broadband screener that has been used since 2014 among all children under seven years old in Korea as a part of the National Health Screening Program (NHSP) for young children [18,20]. The items on the K-DST were categorized in five developmental domains—gross and fine motor movement, cognition, language, socialization, and self-help—and were designed for periodic administration [18]. Although the K-DST uses a zero-to-three-point scale rather than the zero-to-two-point scale more common for other screening tests, the vast number of its 335 items is sufficient to reflect the structural formats and characteristics of other broadband developmental screening tests.

Step I: Delphi survey

The initial questionnaire was developed through three steps (Fig. 1). In the first step, the authors reviewed the report by Eun [20], executed by the Korea Centers for Disease Control and Prevention, which analyzed difficulties in the administration of the K-DST. The report contains two survey results investigating item adequacy of the K-DST: one among 415 parents of young children, and the other for 83 experts [20]. The survey with parents collected opinions about overall problems, while the survey of experts asked for each item. In the
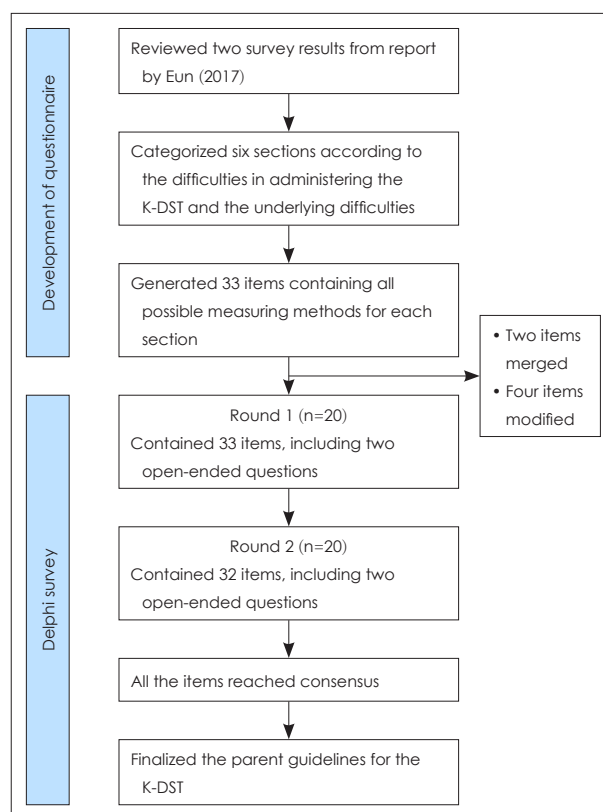

Fig. 1. Flowchart for the development of the questionnaire and the Delphi survey procedure. K-DST, Korean Developmental Screening Test for Infants and Children.

second step, the authors categorized items identified as difficult to understand or confusing into six categories: 1) items that can be administered impromptu assessment; 2) items that cannot be administered impromptu assessment; 3) items that present numbers in the performance criteria, such as "five words" or "ten steps"; 4) items that cannot be administered due to absence of the task tools, absence of opportunities, safety concerns, or other reasons; 5) items that are difficult to understand or confusing; and 6) others. Items in the first and second categories were further classified by whether or not prior observation of the child's behaviors had been made. The sixth category contained items such as the person who administered the test, the location where the test was held, and the length of time spent on test administration. The final step generated a list of items for the questionnaire, containing all possible measuring methods for each category, based on the following sources: 1) user's guidelines for other developmental confirmatory tests [25,26], 2) clinical experience of developmental assessment professionals, and 3) actual experience of parents. The first-round questionnaire comprised 33 items, including two open-ended questions. All items included a comment box (Table 1).

We built an expert panel of 20 experts [27] who had more than 10 years of clinical experience and expertise in pediatric psychiatry, pediatrics, child health nursing, developmental assessment, and special education. Of the 24 experts we initially contacted through email, 20 agreed to participate in the Delphi survey.

We distributed the first-round survey to the expert panel by email. The items in the survey were rated from "very inaccurate" to "very accurate" on a 5-point Likert scale. We calculated the first and third quartiles, the median agreement value (interquartile range=Q_3-Q_1), the convergence value ($X=\frac{Q_3-Q_1}{2}$), and the stability value (COV=$\frac{\sigma}{\mu}$) for each item. The median agreement value 1 or less, the convergence value 0.5 or less, and the stability value 0.5 or less were considered to have reached consensus. If more than two items had the same median value in the same category, then we selected an item with a higher mean value. The additional comments were qualitatively analyzed and reflected in the modification of the items. Results from the previous survey (the ranges of the first and third quartiles and the median value) and the modified items were highlighted in the next round's questionnaire. After all items reached consensus, we selected the items with the highest accuracy score in each category and created the final version of the parent guidelines.

Step II: subjective usefulness survey for the parent guidelines

We surveyed the parents' subjective usefulness to investigate the clinical feasibility of the developed guidelines. Parents aged 18 years or older who had performed the K-DST within the past six months were recruited through the online communities for parents, and duplicate participation was prevented by requiring ID authentication. We targeted a total of 167 parents, calculated as 15 times the 10 influential factors—age, sex, living area, education level, number of children, age of the child, primary caregiver, location where the test was administered, observation of the child prior to administering the test, and scoring methods—with 10% of attrition rate [28]. In addition to the subjective usefulness of each bullet point in the guidelines, we asked about parents' socio-demographic factors and their usual measuring methods of administering the K-DST within the six categories. We calculated the number of subjects and percentages to descriptively analyze characteristics of the participants and the results of the survey. Comments were categorized by the characteristics of the contents and then qualitatively analyzed.

Microsoft Excel® software (2019; Microsoft Corp., Redmond, WA, USA) was used for statistical analysis.

Ethics statement

This study was conducted after approval of the Institutional Review Board of Seoul National University (IRB No. 2007/002-002).

We obtained written informed consent from the expert panels for Delphi survey and online informed consent from the parent participants for online survey prior to their participation.

RESULTS

Delphi consensus on parent guidelines

The expert panel consisted of 20 panelists, four from each specialized field, and they all participated in both rounds. Of the 33 items, including the two open-ended questions, 14 items reached consensus during the first round. Based on the results and the comments from round one, we merged two items and modified four items. Therefore, 32 items were included in the second-round questionnaire, and they all reached consensus after round two. We selected the items with the highest accuracy score in the six categories and modified them for better readability and understandability for parents (Fig. 2). The finalized parent guidelines were approved by the panelists.

Subjective usefulness of the parent guidelines

A total of 167 parents of young children participated in the online survey that investigated the subjective usefulness of the developed guidelines. Among the participants, 132 (80.5%) were in their thirties, 157 (94.0%) were mothers, and 127 (76.0%)

Table 1. Results of the Delphi survey rounds 1 and 2 (n=20)

Items of the Delphi survey	Round 1				Round 2			
	Median*	Agreement	Convergence	Stability	Median*	Agreement	Convergence	Stability
I. Items that can be administered by impromptu assessment								
- When prior observation of the child's behaviors had been made								
1. Score the child's performance according to the number of executions[†]	4.0	1.00	0.50	0.28	4.0[‖]	0.25	0.13	0.12
2. Score the child's performance according to their proficiency, regardless of the number of executions	3.5	1.25	0.63	0.30	4.0	1.00	0.50	0.20
- When prior observation of the child's behaviors had not been made								
3. Administer impromptu assessment and score the child's performance according to the number of executions	3.0	1.25	0.63	0.30	3.0	1.00	0.50	0.14
4. Administer impromptu assessment and score the child's performance according to the number of trials prior to the first successful execution	3.5	2.00	1.00	0.29	4.0	1.00	0.50	0.23
5. Administer impromptu assessment and score the child's performance according to their proficiency regardless of the number of executions	4.0	1.25	0.63	0.30	4.0[‖]	1.00	0.50	0.13
6. Score the child's performance according to different performance previously observed	2.0	0.00	0.00	0.27	2.0	0.00	0.00	0.14
7. Score the child's performance as 0 points	2.0	1.25	0.63	0.55	2.0	1.00	0.50	0.29
8. Leave the items blank[‡]	2.0	2.00	1.00	0.57	4.0	1.00	0.50	0.14
II. Items that cannot be administered by impromptu assessment								
- When prior observation of the child's behaviors had been made								
9. Appropriate length of time observing the child (days)	7.0	7.00	3.50	0.64	7.0	7.00	3.50	0.35
10. Score the child's performance according to the number of executions	3.0	1.00	0.50	0.27	3.0	1.00	0.50	0.14
11. Score the child's performance according to their proficiency, regardless of the number of executions	3.5	1.00	0.50	0.29	4.0	0.25	0.13	0.12
- When prior observation of the child's behaviors had not been made								
12. Score by the assessment of a third party who frequently observes the child with the same standard	3.5	1.00	0.50	0.26	4.0	1.00	0.50	0.14
13. Score the child's performance according to different performance previously observed	2.0	1.00	0.50	0.37	2.0	1.00	0.50	0.20
14. Score the child's performance as 0 points	2.0	2.00	1.00	0.52	2.0	0.25	0.13	0.28
15. Leave the items blank[‡]	2.0	2.00	1.00	0.52	3.0	1.00	0.50	0.17
III. Items that present numbers in the performance criteria								
16. Score the child's performance when performed above the numeric standard as successful	4.0	1.25	0.63	0.24	4.0	0.00	0.00	0.11
17. Score the child's performance when performed near, but below, the numeric standard as successful	3.0	1.00	0.50	0.32	3.0	0.25	0.13	0.16
18. Score the child's performance according to their proficiency, regardless of meeting the numeric standard	3.0	2.00	1.00	0.32	3.0	0.00	0.00	0.21

Table 1. Results of the Delphi survey rounds 1 and 2 (n=20) (continued)

Items of the Delphi survey	Round 1				Round 2			
	Median*	Agreement	Convergence	Stability	Median*	Agreement	Convergence	Stability
IV. Items that cannot be administered due to absence of the task tools or absence of opportunities for safety concerns, or other reasons								
19. Score the child's performance using similar, but different, tools, based on subjective judgement	3.0	2.00	1.00	0.32	3.0	1.00	0.50	0.21
20. Score the child's performance according to different performance previously observed	3.0	1.00	0.50	0.30	3.0	0.25	0.13	0.16
21. Score the child's performance as 0 points	2.0	1.25	0.63	0.55	2.0	1.00	0.50	0.33
22. Leave the items blank‡	2.0	1.50	0.75	0.57	3.5	1.00	0.50	0.20
V. Items that are difficult to understand or are confusing								
23. Assess in consultation with a physician§	4.0	1.25	0.63	0.22	3.5	1.00	0.50	0.20
24. Score the child's performance when performed exactly the same as the item, regardless of understanding of its intention	2.0	1.00	0.50	0.39	2.0	1.00	0.50	0.20
25. Score the child's performance when performed similarly to the item, regardless of understanding of its intention	3.0	2.00	1.00	0.37	3.0	1.00	0.50	0.24
26. Score the child's performance as 0 points	2.0	1.00	0.50	0.55	2.0	1.00	0.50	0.29
27. Leave the items blank‡	2.0	2.00	1.00	0.52	-	-	-	-
VI. Others								
- Person who administered the test								
28. Primary caregiver	4.5	1.00	0.50	0.17	5.0	1.00	0.50	0.13
29. Others	3.0	1.00	0.50	0.25	3.0	1.00	0.50	0.20
- Place where the test was held								
30. At home, prior to the appointment	4.0	1.00	0.50	0.14	4.0	0.00	0.00	0.10
31. At other place, prior to the appointment	3.0	1.00	0.50	0.29	3.0	1.00	0.50	0.16
32. At the clinic in the waiting room	3.0	1.25	0.63	0.26	3.0	1.00	0.50	0.19
- Time for test administration (min)								
33. Appropriate length of time for test administration	20.0	7.50	3.75	0.52	20.0	0.00	0.00	0.18

*accuracy of the method on a 5-point Likert scale; †for example, one point for one successful execution and two points for two successful executions; ‡changed the item to "Leave the items blank and assess in consultation with physician" for the round-two survey; §this was merged with item number 27 in round two; ‖this has the biggest mean value
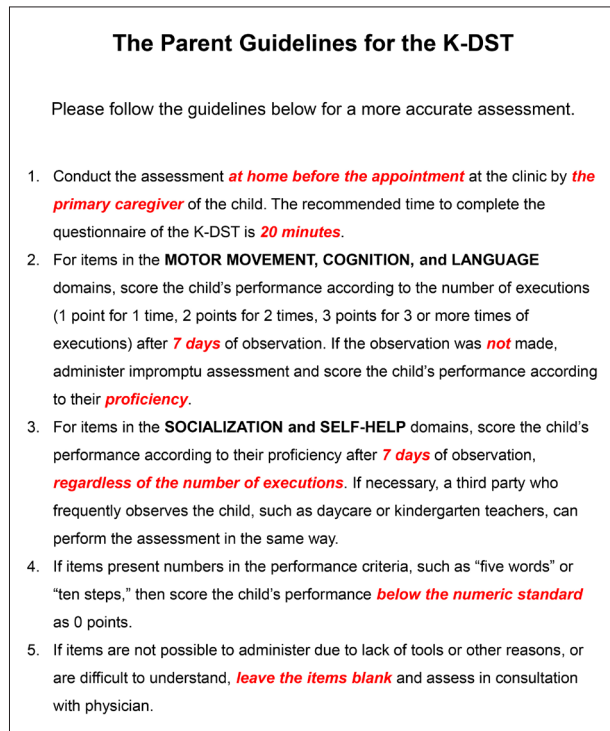
The Parent Guidelines for the K-DST

Please follow the guidelines below for a more accurate assessment.

1. Conduct the assessment *at home before the appointment* at the clinic by *the primary caregiver* of the child. The recommended time to complete the questionnaire of the K-DST is *20 minutes*.
2. For items in the **MOTOR MOVEMENT, COGNITION, and LANGUAGE** domains, score the child's performance according to the number of executions (1 point for 1 time, 2 points for 2 times, 3 points for 3 or more times of executions) after *7 days* of observation. If the observation was *not* made, administer impromptu assessment and score the child's performance according to their *proficiency*.
3. For items in the **SOCIALIZATION and SELF-HELP** domains, score the child's performance according to their proficiency after *7 days* of observation, *regardless of the number of executions*. If necessary, a third party who frequently observes the child, such as daycare or kindergarten teachers, can perform the assessment in the same way.
4. If items present numbers in the performance criteria, such as "five words" or "ten steps," then score the child's performance *below the numeric standard* as 0 points.
5. If items are not possible to administer due to lack of tools or other reasons, or are difficult to understand, *leave the items blank* and assess in consultation with physician.

Fig. 2. Final version of the parent guidelines for the K-DST. K-DST, Korean Developmental Screening Test for Infants and Children.

Table 2. Characteristics of participants of the usefulness survey (n=167)

Characteristics	Value
Age (yr)	
20s	7 (4.3)
30s	132 (80.5)
40s	25 (15.2)
Sex	
Male	10 (6.0)
Female	157 (94.0)
Living area	
Metropolitan area	131 (78.4)
Other areas	36 (21.6)
Education level	
High school graduate or less	18 (10.8)
Bachelor's degree	127 (76.0)
Master's degree or more	22 (13.2)
Number of children	
1	92 (55.1)
2	62 (37.1)
≥3	13 (7.8)
Age of the child assessed	
<36 mo	79 (47.6)
≥36 mo	87 (52.4)

Values are presented as number (%)

had a bachelor's degree (Table 2). When considering the agreements between the parents' usual measurement methods and the guidelines, the majority of participants answered that 1) the primary caregiver of the child administered the test (95.8%), 2) they conducted the assessment at home prior to the appointment at the clinic (77.8%), and 3) they scored the child's performance according to their proficiency for items in the socialization and self-help assessment domains when the observation was made (73.1%) (Table 3). However, only a small number of parents answered that 1) they observed the child for seven days prior to administering the assessment (25, 15.0%), 2) they left items blank and assessed in consultation with physician when the items were difficult to understand (26, 15.6 %), and 3) they scored the child's performance according to the number of executions for items in the motor movement, cognition, and language domains when observations were made (29, 17.4%). Only half of the parents (51.5%) scored 0 points when the child performed below the numeric criteria. For the subjective usefulness of the overall guidelines, 67.7% (113) of the parents thought it was useful. When analyzing more specifically, instructions that recommended a different scoring strategy in answering the questionnaire from their previous measurement tended to be regarded as more useful by the parents. The instructions for items difficult to score due to a lack of assessment tools or inadequate understanding that had around 20% of the agreement between the guidelines and the parents' pre-

vious measurement showed the biggest proportion of the parents (64.1%) answering "useful." The results also revealed a huge gap between the number of parents who answered "useful" and those who answered "unuseful" for each bullet point, ranging from four to nine times. Additional comments were categorized by satisfaction about the guidelines (101, 60.5%), K-DST/NHSP-related comments (50, 30.0%), and others (16, 9.6%) (Supplementary Table 1 in the online-only Data Supplement). Of the 102 satisfaction-related comments, specific standards for test administration (32, 31.4%) was the most frequently mentioned reason in both the satisfactory and unsatisfactory responses. Other parents were satisfied with the simplicity of the guidelines, information about the length of the observation period, and an option for answering the question in consultation with a doctor. A strict scoring method, such as scoring 0 points for a child's performance below the standard, and insufficient simplicity of the guidelines were reasons for dissatisfaction. Among the comments about the K-DST or the NHSP, 78% were related to the difficulties in administering the test.

DISCUSSION

The purpose of this study was to develop guidelines that

Table 3. Number of parents using the same methods suggested in the developed guidelines, grouped by their subjective usefulness (n=167)

Contents of the guidelines	Parents using the same method	Subjective usefulness		
		Useful*	Moderate	Unuseful†
Items in the motor movement, cognition, and language domains		94 (56.3)	51 (30.5)	22 (13.2)
If observation was made, score the child's performance according to the number of executions.	29 (17.4)	17 (58.6)	9 (31.0)	3 (10.3)
If observation was not made, administer impromptu assessment and score the child's performance according to their proficiency.	68 (40.7)	42 (61.8)	16 (23.5)	10 (14.7)
Items in the socialization and self-help assessment domains		84 (50.3)	70 (41.9)	13 (7.8)
If observation was made, score the child's performance according to their proficiency, regardless of the number of executions.	122 (73.1)	64 (52.5)	52 (42.6)	6 (4.9)
If observation was not made, score by the assessment of a third party who frequently observes the child, with the same standard.‡	67 (40.1)	35 (52.2)	27 (40.3)	5 (7.5)
Items presenting numbers in the performance criteria		104 (62.3)	50 (29.9)	13 (7.8)
Score the child's performance below the numeric standard as 0 points.	86 (51.5)	51 (59.3)	34 (39.5)	1 (1.2)
Items difficult to score		107 (64.1)	48 (28.7)	12 (7.2)
If items were not possible to administer due to lack of tools or other reasons, leave the items blank and assess in consultation with a physician.	36 (21.6)	22 (61.1)	14 (38.9)	0 (0.0)
If items were difficult to understand, leave the items blank and assess in consultation with physician.	26 (15.6)	12 (46.2)	14 (53.8)	0 (0.0)
Others		85 (50.9)	68 (40.7)	14 (8.4)
Administer the assessment by the primary caregiver.	160 (95.8)	79 (49.4)	67 (41.9)	14 (8.8)
Conduct the assessment at home before the appointment at the clinic.	130 (77.8)	70 (53.8)	55 (42.3)	5 (3.8)
Observe the child for seven days before administering the assessment.	25 (15.0)	15 (60.0)	7 (28.0)	3 (12.0)
The whole guideline		113 (67.7)	44 (26.3)	10 (6.0)

Values are presented as number (%). *number of subjects who marked "useful" or "very useful"; †number of subjects who marked "unuseful" or "very unuseful"; ‡"If observation was not made, administer impromptu assessment and score the child's performance according to their proficiency" was used instead when asking for the parents' performance

are easy to understand but also comprehensive enough to cover the difficulties parents face when administering the screening tests. The final version of the guidelines provided sufficient information about answering the questions of the screening tests, from "who" and "where" to administer the tests to "how" to score the child's performance on the zero-to-three-point scale. Allowing parents to answer the questions based on objective evidence as much as possible was the priority of these guidelines, and so specific instructions are stated according to the characteristics of the items and by the developmental domains the items measure for. Although these guidelines were developed based on the K-DST, the usability can be expanded to other screeners because of the structural formats and the characteristics of the items in the K-DST.

These instructions exhibited high similarity to the guidelines for the Bayley Scales of Infant and Toddler Development (BSITD), which is widely used as a "gold standard" test to measure developmental status. According to the administration manual of the BSITD [29], each of the 326 items includes a detailed instruction for administration, in addition to the general instructions for length of administration time, number of trials, time measuring method, and so on. A pink three-piece jigsaw puzzle, for example, includes the guidance to provide only one opportunity with a time limit of 180 seconds, which measures child's performance by proficiency. In another case, the BSITD offers guidance for items in which numbers are included in the performance criteria—to answer "yes" only when a child performs the exact number of the criteria. These measurements were also included in the developed guidelines. Although the scoring criteria for the BSITD are different from those of our guidelines—the BSITD includes only yes-or-no questions—most of the contents in our guidelines give instructions similar to those of the BSITD.

The rates of agreement between the guidelines and the parents' usual measurements showed clear deviations based on the characteristics of the measurements, and they reflected challenges to accurately administering the screening tests. According to the survey results, measurements using subjec-

tive evidence, such as scoring the child's performance based on proficiency, had higher agreement rates than those using objective evidence, such as scoring the performance based on the number of successful executions. In the case of items that include numbers, half of the parents gave points when the child did not satisfy the exact numbers in the items, and 5% of those who wrote additional comments about satisfaction of the guidelines thought that scoring 0 points for a child's performance below the standard was too strict. This may lead to an overestimation of a child's abilities and an increase in false-negatives. In fact, a previous study revealed that clinicians lacked trust in parents as administrators of developmental screening tests due to parents' overestimation of their child's abilities and their inadequate knowledge of development [30].

The strength of these guidelines is that they provide instructions based on the characteristics of the questions, so it can be generally applied to other developmental screening tests. It is simple to understand and easy to practice, and it can be used by parents when administering a developmental screening test with their children in a clinical setting. The usefulness survey results exhibited a pattern of the measurements with lower agreement rates showing higher percentages of parents who answered "useful." The measurements for items difficult to score—due to not understanding the meaning or lacking the opportunity to administer, for various reasons, for example— yielded the highest percentage of parents who considered these instructions useful, while about 80% of parents appeared to answer to these items without consultation with a physician. Additional comments from the survey showed high satisfaction of the usefulness of the guidelines, with the highest percentage being the more specific instructions for test administration, which indicates that these guidelines met their needs. Overall, the percentage of parents who answered "useful" for each of the instructions and for the whole guidelines outweighed the percentage of those who answered "unuseful."

As a limitation of this study, the study participants have limited generalizability. Despite the typical size of the Delphi panel group [31], only four experts in each of the five expertise areas were included in the group, which may limit the representativeness of each field. For the usefulness survey, participants were recruited from online communities, so they may not represent the characteristics of the target population.

CONCLUSION

To the best of our knowledge, this is the first study to develop parent guidelines for the administration, not the interpretation, of screening tests, and the first study to use the Delphi technique to develop the guidelines. Findings from the usefulness survey reflected the parents' needs for more spe-

cific scoring standards. Further studies are needed to evaluate the effectiveness of the guidelines in terms of the accuracy of test administration of parents and the diagnostic accuracy of the tests.

Supplementary Materials

The online-only Data Supplement is available with this article at https://doi.org/10.5765/jkacap.230002.

ORCID iDs

Sung Sil Rah   https://orcid.org/0000-0002-3666-4680
Soon-Beom Hong   https://orcid.org/0000-0003-1030-0763
Ju Young Yoon   https://orcid.org/0000-0003-3944-0663

REFERENCES

1) **Marquis S, McGrail K, Hayes M, Tasker S.** Estimating the prevalence of children who have a developmental disability and live in the province of British Columbia. J Dev Disabil 2018;23:46-56.

2) **Rah SS, Hong SB, Yoon JY.** Prevalence and incidence of developmental disorders in Korea: a nationwide population-based study. J Autism Dev Disord 2020;50:4504-4511.

3) **Zablotsky B, Black LI, Maenner MJ, Schieve LA, Danielson ML, Bitsko RH, et al.** Prevalence and trends of developmental disabilities among children in the United States: 2009-2017. Pediatrics 2019;144:e20190811.

4) **Odom SL, Horner RH, Snell ME, Blacher JB.** Handbook of developmental disabilities. New York: Guilford Press;2007.

5) **Dawson G.** Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. Dev Psychopathol 2008; 20:775-803.

6) **Zwaigenbaum L, Bauman ML, Choueiri R, Kasari C, Carter A, Granpeesheh D, et al.** Early intervention for children with autism spectrum disorder under 3 years of age: recommendations for practice and research. Pediatrics 2015;136(Suppl 1):S60-S81.

7) **Harrison M, Jones P, Sharif I, Di Guglielmo MD.** General pediatrician-staffed behavioral/developmental access clinic decreases time to evaluation of early childhood developmental disorders. J Dev Behav Pediatr 2017;38:353-357.

8) **Huttenlocher PR.** Synaptic density in human frontal cortex - developmental changes and effects of aging. Brain Res 1979;163: 195-205.

9) **Barger B, Rice C, Wolf R, Roach A.** Better together: developmen-

tal screening and monitoring best identify children who need early intervention. Disabil Health J 2018;11:420-426.

10) **Agarwal PK, Xie H, Sathyapalan Rema AS, Rajadurai VS, Lim SB, Meaney M, et al.** Evaluation of the ages and stages questionnaire (ASQ 3) as a developmental screener at 9, 18, and 24 months. Early Hum Dev 2020;147:105081.

11) **Glascoe FP.** Early detection of developmental and behavioral problems. Pediatr Rev 2020;21:272-280.

12) **Squires J, Bricker DD.** Ages & stages questionnaires, 3rd ed. Baltimore, MD: Paul H. Brookes Publishing Co., Inc.;2009.

13) **Council on Children With Disabilities; Section on Developmental Behavioral Pediatrics; Bright Futures Steering Committee; Medical Home Initiatives for Children With Special Needs Project Advisory Committee.** Identifying infants and young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. Pediatrics 2006;118:405-420.

14) **Paul H. Brookes Publishing Co., Inc.** 4 decades of development [Internet]. Baltimore, MD: Paul H. Brookes Publishing Co., Inc. [cited 2022 Apr 5]. Available from URL: https://agesandstages.com/about-asq/asq-development/.

15) **Small JW, Hix-Small H, Vargas-Baron E, Marks KP.** Comparative use of the ages and stages questionnaires in low- and middle-income countries. Dev Med Child Neurol 2019;61:431-443.

16) **Paul H. Brookes Publishing Co., Inc.** Translations of ASQ [Internet]. Baltimore, MD: Paul H. Brookes Publishing Co., Inc. [cited 2022 Jul 18]. Available from URL: https://agesandstages.com/products-pricing/languages/.

17) **Sheldrick RC, Marakovitz S, Garfinkel D, Carter AS, Perrin EC.** Comparative accuracy of developmental screening questionnaires. JAMA Pediatr 2020;174:366-374.

18) **Rah SS, Hong SB, Yoon JY.** Screening effects of the national health screening program on developmental disorders. J Autism Dev Disord 2021;51:2461-2474.

19) **Sices L, Stancin T, Kirchner L, Bauchner H.** PEDS and ASQ developmental screening tests may not identify the same children. Pediatrics 2009;124:e640-e647.

20) **Eun B.** Standardization and validity reevaluation of the Korean developmental screening test for infants and children. Cheongju: Korea Centers for Disease Control and Prevention;2017.

21) **Eun B.** Korean developmental screening test for infants and children user's guide. Cheongju, Seoul: Korea Disease Control and Prevention Agency, Korean Pediatric Society;2017.

22) **Glascoe FP, Marks KP, Poon JK, Macias MM.** Identifying and addressing developmental-behavioral problems. Moorabbin, VIC: Hawker Brownlow Education;2016.

23) **Squires T, Bricker P.** ASQ for parents [Internet]. Baltimore, MD: Paul H. Brookes Publishing Co., Inc. [cited 2022 Jul 18]. Available from URL: https://agesandstages.com/wp-content/uploads/2018/12/ASQ-For-Parents-Packet.pdf.

24) **The Royal Children's Hospital Melbourne, Centre for Community Child Health.** PEDS brief administration and scoring guide [Internet]. Parkville, VIC: RCH [cited 2022 Jul 26]. Available from URL: https://www.rch.org.au/uploadedFiles/Main/Content/ccch/PEDS-Brief-Administration-and-Scoring-Guide.pdf.

25) **Bayley N.** Bayley scales of infant and toddler development, Third edition: administration manual. San Antonio, TX: Pearson Psychcorp;2006.

26) **Raiford SE, Coalson DL.** Essentials of WPPSI-IV assessment. Hoboken, NJ: John Wiley & Sons;2014.

27) **Jorm AF.** Using the Delphi expert consensus method in mental health research. Aust N Z J Psychiatry 2015;49:887-897.

28) **Pett MA, Lackey NR, Sullivan JJ.** Making sense of factor analysis: the use of factor analysis for instrument development in health care research. Thousand Oaks, CA: Sage Publications;2003.

29) **Bayley N.** Bayley scales of infant and toddler development, Third edition (Bayley®-III). San Antonio, TX: Pearson Psychcorp;2006.

30) **Morelli DL, Pati S, Butler A, Blum NJ, Gerdes M, Pinto-Martin J, et al.** Challenges to implementation of developmental screening in urban primary care: a mixed methods study. BMC Pediatr 2014;14:16.

31) **Ali N, Rigney G, Weiss SK, Brown CA, Constantin E, Godbout R, et al.** Optimizing an eHealth insomnia intervention for children with neurodevelopmental disorders: a Delphi study. Sleep Health 2018;4:224-234.