VALIDITY AND REALIBILITY TESTING OF THE ALPHABET KNOWLEDGE MEASUREMENT INSTRUMENT

1Angela Mahkota, 2Denny Putra
1Universitas Kristen Krida Wacana
mahkotaangel@gmail.com
2Universitas Kristen Krida Wacana
denny.putra@ukrida.ac.id

Abstract

This article focuses on the early childhood education program (PAUD) in Kelurahan Tanjung Duren Utara. The aim of this study is to examine the validity and reliability of a measure of alphabet knowledge among children aged five to six years. The hypothesis of this research is that the measure of alphabet knowledge is tested to be valid and reliable. This research adopts a quantitative approach and the total sampling method. Participants’ alphabet knowledge is measured with an instrument developed by the researcher based on definitions and factors related to alphabetical knowledge theories. The validity test showed the measure of alphabet knowledge to be valid (0.83-1 > 0.5) and reliable (0.94> 0.9). This research examines 78 items that are valid and reliable for assessing alphabet knowledge among children aged five to six. Further findings based on an analysis of item difficulty evidenced that the letters v, q, p, f among others are difficult to learn. Item discrimination was also analyzed, and results revealed that for certain vowels it was difficult to discriminate between a child with and without alphabet knowledge. A distractor power analysis of alphabet knowledge on the other showed two items in particular that need to be revised, namely letters N and t.

Key words: Early Childhood, Alphabet Knowledge, Validity, Reliability.

INTRODUCTION

Early childhood education (PAUD, pendidikan anak usia dini) helps prepare children before entering primary school (Arijani, 2013). Primary school preparations include knowledge of the alphabet, writing and counting (Halimah and Kawuryan, 2010). Education and Culture Ministerial Regulation No. 134/2014 stipulates that the early literacy skills of children aged 5-6 should cover initial familiarity with the first letters of objects around them, the ability to make the connection between sounds and letters, writing and reading their own name, and understanding the meaning of words in a story. According to Santrock (2009), preschool-aged children are ready for school and behave more independently, and have begun learning to follow instructions and identify letters or generally known as alphabet knowledge. Piasta & Wagner (2010) stated that alphabet knowledge relates to the knowledge of the names and sounds of letters.

Preschoolers should be taught the ability to recognize letters as an early learning process before they can read fluently. According to Cardenas, Furlong & Garcia (2016), alphabet knowledge is closely associated with the ability to form syllables in the early stage of reading. Based on an observation of a PAUD center located in Kelurahan Tanjung Duren Utara, 22 out of 24 children are weak in alphabet knowledge. The researcher became aware of the problem when asking questions on letters and the children were unable to answer correctly. The researcher also found that these children could write the alphabet but did not know the names to the written letters.

This is also a problem in other regions throughout Indonesia as observed by several studies that showed weak alphabet knowledge due to inability to recognize letters and distinguish among them (Sundari & Masudah, 2014; Rohmawati & Khotimah, 2015; Pebriani, 2015). This issue has thus far been examined only through observations that do not provide a definite measure of alphabet knowledge. As studies in Indonesia are in fact interventions to address alphabet-knowledge issues, they are focused on improving alphabet knowledge. From these interventions, it was found that letter recognition has
improved by using various methods such as the use of decorative bottle caps, sandpaper letters, bowling and visual aids (Muflikha, 2013; Candra, Djaelani, & Rahmawati, 2013; Aryani, Agung, & Tirtayani, 2015; Wahdaniah, Fadilah, & Yuniarni, 2014). The researcher conducted a literature review but did not find any instrument that specifically measures alphabet knowledge in Indonesia except for the emergent literacy test.

The researcher learnt that the emergent literacy test was developed by an international organization based on the theory posited by Whitehurst & Lonigan (1998). One of the elements of emergent literacy is alphabet knowledge. The two emergent literacy screening tools are known as PALS-PreK (Phonological Awareness Literacy Screening) dan Get Ready to Read. The researcher chooses not to use these tools as they do not specifically measure alphabet knowledge. The researcher came across another organization at www.lakeshorelearning.com that has developed an instrument to specifically measure alphabet knowledge. The instrument however has no clear theoretical basis and has not been translated into the Indonesian language, thus the researcher believes that there are differences in the level of difficulty for certain letters, in addition to differences in settings in regard to the learning process where children are taught to spell in the Indonesian language.

In view of this, the researcher is interested in developing an instrument to measure alphabet knowledge that covers the ability to recognize the names and sounds of letters, and to write them down. The researcher therefore shall test the validity and reliability of the instrument on children aged 5-6. The research problem is on whether the instrument developed by the researcher to measure the knowledge of the alphabet in the Indonesian language is both valid and reliable. Research results are expected to provide insights on the construction of the alphabet knowledge measurement instrument in relation to the validity and reliability testing as not many instruments are available for assessing alphabet knowledge. This research also hopes to contribute to the community of schools, teachers and parents, especially in Kelurahan Tanjung Duren Utara, Jakarta, in making the instrument available for assessing the alphabet knowledge of children aged 5-6.

Alphabet Knowledge

Alphabet knowledge refers to the attachment of names to letters (NELP, 2008). According to Whitehurts & Lonigan (1998), alphabet knowledge is a child’s ability to recognize the names of letters by associating sounds to printed letters, and decoding written letters into sounds. Alphabet knowledge also includes the ability to write by translating sounds into written words. Worden & Boettcher (1990) presented several critical elements of alphabet knowledge, i.e., letter recitation, naming and printing. The alphabet consists of the uppercase and lowercase letters, and is also grouped into vowels and consonants. According to the Indonesian Language Development Board (2016), vowels consist of the letters a, i, u, e, o. The letter “e” can be pronounced differently as “è” and “é”. In Indonesian, consonants comprise of the letters b, c, d, f, g, h, j, k, l, m, n, p, q, r, s, t, u, v, w, x, y, z.

Factors Affecting Alphabet Learning

The four factors that can affect the learning of the alphabet, according to Treiman (in Jones, Clarrck, & Reutzel, 2012) are the position of the letters, letter-names that contain sounds, shape or form of letters, and the differential rates of exposure to letters (in Jones, Clarrck, & Reutzel, 2012). Letters located at the beginning and end of the alphabet are much easier to learn than letters in the middle of the alphabet. For example, the letter “a” at the beginning and “z” at the end of the alphabet. Letter-names that contain the sound represented by the letter are learned more easily than letter-names where the sound is not represented by the letter. For example, the letters b/ bi, d/ di in English can be learned with greater ease compared to h/ eich, w/ dobliyu.

Letters with distinctive shape or form are much easier to learn than letters having similar shape or form. The letters b and d, as well as O and Q are more difficult to differentiate than the letters c and h, k and l. Differential rates of exposure to letters, such as letters found in a child’s own name or letters that appear frequently in printed reading materials in the child’s environment can also help the child learn and understand letters better. According to Wood & Mclemore (2001), children can more easily visualize capital letters than small letters, therefore the uppercase form is normally taught earlier than lowercase letters.

Santrock (2009) defines early childhood as a time that spans from three to five or six years of age, or typically known as preschool age. Based on Piaget’s theory (in Santrock, 2009), the concrete preoperational stage comprises of the symbolic function and intuitive thought sub-stages. Children aged 2 to 4 go through the symbolic function sub-
stage where they develop the ability to think about objects that are not present or real, and their symbolic thoughts heightened. Regarding early childhood development, Education and Culture Ministerial Regulation No. 134/2014 describes children age 4-5 as already having literacy abilities, such as the ability to identify animal sounds around them, as well as write and say the alphabet from A to Z.

Children aged 5-6 can already recognize the initial letters of objects around them, make the connection between letter-sounds and forms or shapes, write and read their own names, and understand the meaning of words in a story. Preschoolers who often listen to the ABC alphabet song are more likely to recognize with greater ease the letters positioned at the beginning of the alphabet, while cognitive processing skills and phonological awareness in children can affect their ability to acquire alphabet knowledge (Philips, Piasta, Anthony, Lonigan, & Francis, 2012). According to Santrock (2009), the cognitive development of children aged 4-6 already involves the development of schemas from which they can build understanding and classify shapes and forms.

Validity & Reliability

Validity refers to the accuracy of a measurement or testing instrument to gauge the accuracy of the said instrument (Azwar, 2009). Sugiyono (2011) on the other hand stated that a valid measurement instrument can confirm the accuracy between actual data and data reported by the researcher. There are three types of validity: construct validity, content validity and criterion validity. Content validity looks into the extent to which the content of the test or measurement instrument reflects the attributes to be measured to ensure that the content is relevant and remains within the boundary of measurement objectives. Construct validity shows the degree to which the test proves the theoretical construct to be measured. Criterion validity is divided into predictive and concurrent validity. Predictive validity is a test that can function as a predictor of sometime in the future. Concurrent validity refers to the scores for the test and criterion assessed at the same time, and the correlation between both scores.

Reliability is the consistency of a measurement or testing instrument that provides reliable results (Azwar, 2009). Reliability shows consistency over a specified time interval (Sugiyono, 2011). Reliability testing consists of internal and external reliability. External reliability testing can be done through the test-retest method and the equivalent form, or a combination of the two. Internal reliability can be tested through different techniques such as Spearman Brown (split half), KR 20, KR 21, and Anova Hoyt.

Item Analysis

Quantitative item analysis according to Urbina (2007) is divided into three categories: item discrimination, item difficulty and distractor power. Item discrimination analysis looks into how accurately an item is able to discriminate between good and poor test takers in respect to behavior, knowledge or other characteristics that are normally analyzed through the discrimination index (D). Item difficulty analysis examines the level of difficulty of test items, thus these items can be arranged according to the test objectives. Distractor power analysis identifies every distractor of an item that is functioning effectively, which is chosen relatively evenly by test takers.

Issues related to alphabet knowledge have fairly caught the attention of several researchers in Indonesia. Studies found that children in Indonesia still have low-level knowledge of the alphabet (Sundari & Masudah, 2014; Rohmawati & Khotimah, 2015; Pebriani, 2015). In view of this, Indonesian researchers have initiated interventions to improve alphabet knowledge. Unfortunately, however, despite the interventions to address the issue, children’s level of alphabet knowledge is not being specifically measured. The researcher has come across an organization that has developed an instrument based on the emergent literacy theory of Whitehurst & Lonigan (1998). The instruments are called PALS-PreK (Phonological Awareness Literacy Screening) and Get Ready to Read. The researcher chooses not to use either of these instruments because they still do not measure all aspects related to alphabet knowledge, and do not specifically measure alphabet knowledge. Another organization at www.lakeshorelearning.com has developed an instrument to specifically measure alphabet knowledge but lacks a solid theoretical foundation and has not been translated into the Indonesian language.

The researcher also considers factors that can affect validity and reliability in order to ensure that the instrument developed by the researcher is both valid and reliable. Factors affecting validity are the definite construction of items, the right timing for testing, number of items that should not be too few, and instrument instructions (Sukarfi, 2008). Factors that can affect reliability include test length and item difficulty. The longer the test, the more items can be measured, and items that are too easy or too difficult
The researcher uses the complete list of alphabet from A to Z in order to measure the entire range of materials taught. This research studies the content validity of the test, thus requiring professional judgment to assess the instrument that the researcher has developed, which was later analyzed with Aiken's V. For this research, internal reliability was tested to look at the consistency of scores among test items that are analyzed with KR 20. If the instrument is valid and reliable, it can be used on subjects having the same criteria. The research hypothesis is that the alphabetical knowledge measurement instrument is tested to be valid and reliable.

RESEARCH METHOD

This research adopts the quantitative method through validity and reliability testing. Participants are children aged 5-6 from the PAUD center in Kelurahan Tanjung Duren Utara, totaling some 70 preschoolers. The total sampling technique is applied for this research. Participants are asked to answer questions on their knowledge of the alphabet through the instrument which the researcher has developed according to Whitehurst & Lonigan's theory (1998) and factors that affect the acquisition of alphabet knowledge (Jones, Clarck, & Reutzel, 2012). The instrument comprises of 78 items. Answers for the test are either correct (1) or incorrect (0) with a maximum score of 78 and a minimum score of 0. Categories are divided into very low 0-15, low 16-30, medium 31-48, high 49-63, and very high 64-78.

The reliability coefficient measured with KR 20 stands at $\alpha = 0.94$, while test validity analyzed with Aiken’s V ranges between 0.83 and 1. This measure also uses the additional analysis of item difficulty, item discrimination and distractor power. Research preparations are preceded with an observation of the phenomenon on the ground before determining the research topic, which is the alphabet knowledge of preschool-aged children. The researcher then collected field data and theories relevant to the research topic. The researcher formulated the research problem, came up with the research title, and developed the hypothesis. The next step is identifying the research population and sample, and sampling technique, and collecting the necessary data.

The researcher developed an instrument for obtaining data on the alphabet knowledge of preschoolers aged 5-6 in Kelurahan Tanjung Duren Utara. The research design is for the researcher to first find information on the number of PAUD centers located in Tanjung Duren Utara. The researcher prepared a letter of permit for data collection, and solicited the help of a tester for conducting the test. The tester is first provided with the necessary information before conducting the actual test. Test takers are given instructions which they must fully understand. The test is conducted on an individual basis for roughly ten minutes for each child. Test results are then analyzed and discussed.

RESULTS AND DISCUSSION

The number of data processed totaled 75 ($N = 75$). The normality test result is $p = 0.073 > 0.05$. The instrument is valid and reliable with a validity coefficient value of 0.83-1 and a reliability coefficient value of 0.94. In regard to normal distribution, the results ($M = 61.68$, $SD = 13.15$) are considered high. The analysis of item difficulty for letter-name knowledge consists of 24 items classified under the easy to very easy category, i.e., m, c, V, I, x, b, F, w, g, O, u, h, $S$, Q, z, e, D, R, t, p, K, A, l, y, J, while item N falls under the moderate category. Letter-sound knowledge covers 22 items from the easy to very easy, i.e., M, C, I, X, B, n, W, G, j, o, U, H, s, z, E, d, r, T, P, k, a, Y, while the items L and f are considered moderate, and items V and q as difficult. On the writing aspect, items classified as easy to very easy are d, u, k, e, s, b, l, t, o, h, m, r, x, p, c, a, I, while z, w, j, g, y, n are moderately easy, and v, f, q are difficult.

Item discrimination to assess letter-name knowledge consists of items m, V, x, b, N, e, D, R, t, p, y, f, P. Poor items under the discrimination index are c, I, w, g, j, O, u, h, Q, z, K, a, l. Meanwhile, good items for letter-sound knowledge are M, V, X, B, n, f, W, G, j, H, s, q, Z, d, r, T, P, L, Y, E, C, while the poor items are I, o, U, k, a, u. Good items in relation to letter-writing cover d, k, v, e, s, v, l, t, h, z, m, r, w, f, x, p, g, c, y, n, q, j, while the poor items are u, o, a, i. Further analysis is conducted on distractor power for letter-name knowledge. The distractor of two items need to be revised, namely item N with options W
and M, and item t with options f and z.

The validity and reliability of the measure developed by the researcher to assess alphabet knowledge has been tested and proven. The minimum coefficient of the validity test that uses Aiken’s V is > 0.5 (Azwar, 2012). The validity coefficient value of the instrument lies between the 0.83 - 1 range. This shows that the instrument can be counted on as a rational analysis, where professional judgement has been sought in the field of clinical psychology, child education psychology, and psychological measurement by considering factors and theories regarding alphabet knowledge that are to be measured. Experts provided feedback on certain items related to letter-sound knowledge that need to be improved, particularly on the sound pronunciation of the letter “e” due to differences in pronouncing the letter c as cê, and f as ef. Clarity in pronouncing letters is important to ensure the proper sound pronunciation of the letters.

The reliability of a test can reflect the homogeneity or sameness of the items in the test (Azwar, 2009). Based on the reliability test, the instrument for measuring alphabet knowledge has a KR 20 reliability coefficient value of 0.94 > 0.8. This shows that the instrument has high reliability given the consistency in the scores of each item for different aspects in the measurement of alphabet knowledge. The high reliability coefficient value is among others attributed to the assessment of 78 items by the instrument that are based on 3 key aspects, whereby each aspect consists of 26 items of the letters A to Z. According to Sukardi (2008), test length can affect reliability as the longer the test, the more items are measured to assess learning materials. Item difficulty is another factor that can influence reliability. Items that are either too easy or too hard can lower reliability (Sukardi, 2008). Based on analysis results, item difficulty in the instrument ranges from very easy to hard, which means varying levels of difficulty that render the instrument as having high reliability.

An analysis of item difficulty shows the need to change the position of items in the instrument. This is because arranging the items from the easiest to the hardest is necessary to keep subjects motivated (Urbina, 2007). Items that fall under the category of easy to very easy in respect to letter-name knowledge are letters m, c, v, i, x, b, f, w, g, o, u, h, s, q, z, e, d, r, t, p, k, a, Y, while the writing of letters d, u, k, e, s, b, l, t, o, h, m, r, x, p, c, a, I are categorized as easy to very easy. Based on the analysis results, the researcher shall arrange the items from the easiest to the hardest for every aspect being measured.

Concerning factors that affect the learning of the alphabet, Treiman (in Jones, Clarck, & Reutzel, 2012) believes that the position of letters influences a child’s ability to learn the alphabet. Letters located in the beginning and end of the alphabet are much easier for children to recognize. These items contain the letters “a” and “z” that are found in the beginning and end of the alphabet. The letter “q” is considered very difficult in terms of sound knowledge as evidenced from the fact that the lowercase “q” is hard for children to learn. According to Treiman (in Jones, Clarck, & Reutzel, 2012), similarities in the shape of letters can affect a child’s ability to recognize letters. Based on observation, the letter q is mostly recognized as the letter p.

The similar looking letters of p and q are quite confusing for children. Both letters are similar in shape except for the direction of the curved stroke, either to the right or left. Mueller & Weidemann (2012) stated that alphabetic letter identification is influenced by the visual similarities of the shape of letters. This is evident in writing the letter z where errors occur because it faces the opposite direction. Reversible letters that look the same but turned in opposite directions include b and d, or J and L. This can be confusing for children particularly if they do not remember or pay attention to the left-right orientation (Treiman & Kessler, 2011). From studying the progress and development of children between the age of 2 and 7, Santrock (2009) found that they have begun learning to form schemas, which means building their understanding and classification of shapes, sizes and colors. The classification of shapes can be in the form of geometric shapes or symbolic representation of letters.

Based on Piaget's level of child development, symbolic thoughts are also heightened in the preoperational stage (Santrock, 2009). In Ellefson, Treiman, & Kessler’s (2009) study of progress in alphabet learning, children will learn about the letter name-shape relationship through memorization, and later on become familiar with the sound of the letters. Test results showed that preschoolers in Kelurahan Tanjung Duren Utara have no difficulty recognizing the letter p. This signifies that the shape of the letter p is already classified as such in the child’s memory. A
child however may have trouble identifying the letter q because he only remembers the letter p, and as a consequence when showed the letter q, most children identified it as the letter p.

According to Wijanarko & Rully (2015), vowel letters are much easier because those are the first letters that children are taught before they are introduced to consonants. The consonants v and f are considered difficult for children to learn as they tend to pronounce these consonants as the sound of the letter p. Based on the item analysis, the letter v is also difficult in terms of letter-sound knowledge as it has been observed that many children identified it as the letter f. In light of this, the researcher is aware that the sound of letters is another factor that affects difficulty in differentiating between the letters v and f as the similarity in sound also have a bearing on a child’s ability to recognize letters. Articulating letters with “e” sound at the end such as “v” and the “e” sound at the front such as “f” tend to be confusing when trying to remember how it sounds. Similar-sounding letters are much harder than letters having different sounds. (Wickelgren, 1966).

Familiarity with letter sounds will help children better learn the alphabet and later understand the relationship between the alphabet phonetic sounds and word formation (Blevins, 1998). The analysis of item discrimination showed that vowels except for the letter e are poor items. The letter e is a good item as it has its own level of difficulty in which there are different sound pronunciations in the letter. This is evidence that letters which children can easily learn are ineffective in discriminating children who have acquired alphabet knowledge from those who do not.

From the analysis of the distractor power, the options of two items need to be revised. The letter N in item 7, i.e., the distractors W and M, from which many participants have chosen the letter M. Based on observation, this is because a large number of children were confused with the similar sounds of the item. The distractor W on the other hand is easier for children to distinguish as neither the sound nor shape is the same with the letter N. The letters N and M have similar sounds when articulating the letter, and the similarity in shape have confused the children. According to Wickelgren (1996), similar-sounding letters are more difficult than letters that sound differently.

Blevins (1998) also highlighted on the confusing similarities in letter shape, stating that letter shapes with visual similarities such as the letters p-q, M-N, b-d are indeed confusing. Item t number 21 with distractor f and z also need to be revised. The letter f is more likely to confuse children as its shape is similar to the letter t, with both having a cross stroke. From observation during the data collection process, several children mentioned that the horizontal stroke in the middle is the letter “f”. Blevins (1998) also noted that the letters f and t have similar shapes confusing to children. Similarities in shape can mean letters facing different directions, left or right, up or down, or the curved stroke forming a letter.

In terms of normal distribution, preschoolers in Kelurahan Tanjung Duren Utara have high-level knowledge of the alphabet. This research finding differs from the early observation carried out by the researcher. This is due to the time interval of roughly 8 months from when the observation was conducted by the researcher for data collection purposes. Within the span of 8 months, the children have continued to learn and improve. They are taught the same alphabet lessons on a daily basis at the PAUD center. This undoubtedly affects the children’s ability as they go through a period of imbalance every six months of their life, which refers to the environmental factors that drive a child to learn (Hurlock, 1978).

Repetition in learning can improve long-term memory (King, 2010). Preschoolers learning the alphabet can do so at home with their parents. According to Nuraeni (2011), parents play a critical role in supporting the literacy development process, including alphabet knowledge. Robins, Treiman and Rosales (2014) pointed out that literacy activities that include letter recognition can be reinforced when parents help their children to spell and identify letters. Home learning activities where lessons are repeated can help children gain better memory of the alphabet. Parents therefore are influential in a child’s mastery of the alphabet. This high-level of alphabet knowledge is likely due to an erroneous observation on the researcher’s part who may have inadvertently studied preschoolers with poor alphabet knowledge.

A research limitation concerns the considerable time needed for the testing procedure according to the child’s ability, thus the targeted number of participants in a day may not always be met. Another shortcoming relates to the order of items by level of difficulty which may not have been arranged appropriately. A research advantage however is that the researcher’s instrument can provide evidence on which items are difficult and are affected by factors that have a bearing on the ability to learn the alphabet. A research implication to the Indonesian public is the ability to use the instrument as required.
Based on findings, educators can prioritize on teaching easy-to-learn letters first in order to make it easier for children to recognize the alphabet letters. Theoretically, this research implication is similar to other studies that have shed light on the issue of alphabet knowledge but could only provide interventions without definite measures for Indonesia. This research presents a more definite measure of alphabet knowledge among preschool-aged children. It also provides a more clear-cut theoretical foundation by factoring in the child’s development, as well as other dimensions that affect the ability to acquire alphabet knowledge.

CONCLUSION AND SUGGESTION

Based on this research, the instrument is evidenced to be valid and reliable in assessing the alphabet knowledge of children aged between five and six. The instrument covers 78 items that are proven valid at a coefficient value of 0.83-1 as these items are effective in measuring alphabet knowledge according to the existing constructs. The instrument is also confirmed reliable with a coefficient value of 0.94. It therefore can be trusted as a means to measure the ability of children aged 5-6 in learning the alphabet.

Given the shortcomings which the researcher is mindful of, the following recommendations are put forward to fellow researchers sharing similar interest in the topic. Subsequent research can perform another validity test to examine whether the instrument can become a predictor and additional reliability test to further confirm the reliability of the instrument. The instrument can be applied on a broader geographical scope beyond Kelurahan Tanjung Duren Utara and the application of such instrument should consider the purpose of using the instrument. If it is meant to gain insights into students with strong and weak abilities, using good items in the discrimination index should be able to discriminate between the different levels of competency.

The researcher recommends that PAUD Kelurahan Tanjung Duren Utara makes use of the instrument for measuring alphabet knowledge to assess the ability of their students. Subsequent studies need to spend more time for building a rapport with relevant parties to make it easier for the researcher to approach and engage with the institution. The researcher also needs to consider ways for creating an environment conducive for the testing process, such as sitting at the very back of the classroom. This is necessary to avoid disturbing the classroom learning process. The researcher must also bring at least one research assistant along to avoid spending too much time in a given location.

REFERENCES


