Expert Judgment of Learning Achievements Evaluation Instrument for Children Age 4-5 Years Old

Suci Rohmadheny, Yunisa Laila
Pendidikan Guru Pendidikan Anak Usia Dini, Universitas Pahlawan Tuanku Tambussai
DOI: 10.31004/obsesi.v5i1.524

Abstract
This article discusses the results of expert judgment on the process of developing an evaluation instrument for the development of children aged 4-5 years based on the 2013 Early Childhood Education (PAUD) curriculum. The development of the instrument uses the ADDIE model (Analysis, Design, Development, Implementation, and Evaluation), but this research was carried out at the development stage. One of the development processes in this case is content validation through expert judgment. The expert judgment procedure is carried out by early childhood education experts. After the results of expert judgment are obtained, the researcher examines the indicator items that have been developed for improvement. The content validation was carried out on 186 items of child development achievement indicators that contained 6 aspects of the development of children aged 4-5 years and 46 basic competencies that were tied to 4 core competencies. As a result, as many as 186 of the indicator items received a score of 491 from a maximum score of 744 and a number of recommendations for improvement from experts. This means that the instrument is feasible to be continued at the implementation stage, namely the testing of the instrument, after minor repairs are made according to notes from the expert.

Keywords: early childhood; evaluation; development achievements; learning achievements; content validity
INTRODUCTION

The success of learning in early childhood can be seen from the achievement of their development in all aspects of its development. To be able to see these achievements on an ongoing basis, it is necessary to carry out a process of assessment and evaluation of their progress. In the 2013 PAUD curriculum, the success of learning is aimed at completing basic competencies and developing them in the content of learning material that has been adapted to the standards of child development achievements in each age group (Suminah et al., 2015).

Assessment or assessment is defined as the process of obtaining information that is used to make decisions about the extent of student achievement, curriculum, programs, and education policy (Gullo, 2005; Nitko, 1996). It can be said that the assessment needs to be done before evaluating. The data obtained from the valuation results will be used as material for evaluating. An evaluation process cannot occur without data obtained through the results of the assessment.

To be able to carry out an evaluation and evaluation process, both the process and learning outcomes need to pay attention to the principles of assessment, techniques and instruments, mechanisms, and procedures that must be carried out in the assessment. Based on the regulation of the Minister of Education of the Republic of Indonesia regarding the 2013 PAUD curriculum in the fifth annex, it is stated that the assessment process and learning outcomes function for children aged 4-6 years (TK), while for ages from birth to 4 years the assessment serves as a reference considered in implementing growth stimulation. In general, assessment of early childhood functions to monitor the progress of the learning process, learning outcomes, and improvement of the results of learning activities as a whole and continuously (Republik Indonesia, 2014).

The assessment process and results carried out by early childhood educators in Indonesia use an authentic assessment approach. An authentic assessment is an assessment or evaluation that is based on real data collected from the process and results of children's activities during learning (Bagnato, 2007; Neisworth & Bagnato, 2004; Puckett & Black, 2000; Puspita & Susilowati, 2015). Using this authentic assessment approach, an educator needs to pay attention to principles and characteristics in carrying out an authentic assessment.

There are several characteristics of authentic judgment. These characteristics include: there is a lot of data generated through the collection from time to time, all aspects of child development are interesting and need to be evaluated rather than the results/products of children's work, carried out in the context of natural learning, authentic assessment is functional and embedded in the curriculum, based on the disclosure of the child's best performance, useful as a basis for planning learning, and authentic assessments involving various roles of responsibility: teacher, child, parent, and other professionals (Bagnato, 2007; Sani, 2016).

Educators who will carry out an assessment and evaluation of the process and learning outcomes need a valid and reliable instrument. The instrument basically can be developed by the educator himself, adjusted to the needs of assessment. In an evaluation or assessment instrument, indicators are needed that reflect the achievement of basic competencies that lead to the achievement of the standard level of achievement of child development. These indicators of developmental achievements serve as guidelines in assessing and evaluating early childhood, and can be used as an inspiration for educators in designing learning activities that refer to learning objectives, which are indicators of their developmental achievements (Puspita & Susilowati, 2015).

During this time, there are still many educators who do not evaluate the achievements of child development in accordance with real conditions. This is caused by one thing, namely the tool or instrument used is not yet valid and reliable, so it becomes a problem for educators in evaluating the achievements of child development. One of the causes of these problems is that educators have difficulty in developing instruments for
assessing children's development outcomes. Based on the results of previous studies, it is stated that teachers face a variety of problems in evaluating learning in early childhood, one of the problems faced is the difficulty in preparing instruments (Rohmadheny et al., 2019).

Other studies mention the results of mapping the ability of PAUD teachers in conducting assessment activities for evaluation purposes in Pekanbaru City, which are relatively low (Novianti et al., 2013). The teacher's ability to conduct assessments is needed to support the success of the assessment and evaluation process. However, the varied educational background of PAUD teachers currently has a significant relationship with their competencies while on duty (Gazali, 2012).

This is supported by the results of unstructured interviews conducted by researchers to several teachers randomly from several different agencies in the regions of East Java, Yogyakarta, and Central Java. Some teachers do not yet understand and are not skilled at how to formulate development indicators independently based on the standard level of child development achievement and basic competencies set in the 2013 PAUD curriculum. Therefore, this research is a basic research that tries to help facilitate teachers in overcoming difficulties related to the development of learning evaluation instruments.

This research is a series of development research conducted until the early stages of development. In the initial stages of developing this instrument, there is a validation process that must be carried out. Content validation is done through expert judgment. The process was carried out to obtain input into the improvement and feasibility of the instrument before entering into the next research stage, namely the instrument trial.

**METHOD**

This development research uses the ADDIE model (Analysis, Design, Development, Implementation, and Evaluation), but this research was carried out at the development stage. One of the development processes, in this case, is to develop instruments based on the results of the instrument grating design and validate the contents through expert judgment. This content validation process is presented and discussed in the article. The expert judgment procedure on the validated instrument is still the recommended method because it is considered as an alternative method which is more satisfying than the calculation using the tool (Hughes, 1996; Lad & Kulkarni, 2010; Pulkkinen & Holmberg, 1997; Werner et al., 2017). The following is an expert judgment procedure chart.

![Chart 1. Expert Judgment Procedure for Learning Evaluation Instruments Children aged 4-5 years](image)
This method is carried out by involving experts in relevant fields, namely experts in early childhood education. After this procedure is completed, the researcher does the saving again and fixes the instrument according to the validator’s notes. The validation sheet is presented with instructions for the validator, validation assessment criteria from a scale of 1-4, and a validation instrument table that contains basic competencies, learning objectives, indicators of child development achievements, score fields, and validator notes.

The instrument that gained expert judgment was developed based on the standard level of child development achievement limited to 4-5 years of age, as stipulated in the Regulation of the Minister of Education and Culture of the Republic of Indonesia Number 137 of 2014. This standard is the final goal of a learning program in PAUD. These child development achievement standards are then linked to basic competencies so they can be used as a reference in developing instruments for evaluating child development achievements. The following shows the validation sheet.

Figure 1. View of Validation Instrument for Evaluating the Development Outcomes of Children Aged 4 - 5 Years

The validation sheet above is completed with the results of the validation assessment criteria. The instrument gets a decent criterion when the total validation score is 558 - 744, whereas when the score is in the range of 372 - 557, the instrument is in the eligibility criteria with minor improvements. When the instruments get a total score in the range 187 - 371, then the instrument is in the criteria worthy of major improvement. The lowest is when the total score is in the range 0 - 186, the instrument obtains criteria not suitable for use. These criteria are the result of calculating the total score on a scale of 1-4 for all indicator items. More clearly, the criteria are displayed as follows.

Table 1. Content Validation Assessment Criteria

| Range of Total Score | Criteria               |
|----------------------|------------------------|
| 558 – 744            | Feasible               |
| 372 – 557            | Feasible with Minor Revision |
| 187 – 371            | Feasible with Major Revision |
| 0 – 186              | Not Feasible           |
RESULTS AND DISCUSSION

After an expert assessment of 189 indicator items has been completed, it can be seen the results of the validation of the assessment instruments for the development of children aged 4 - 5 years based on the 2013 PAUD curriculum obtained a score of 491 and some notes from the validator. Get 39 indicator items given a special note by the validator. This note is the main concern of researchers to make improvements to the instrument.

The score procuration in each variable indicator item was explained as follows. A total of 149 indicator items received a score of 3, as many as 16 indicator items scored 2, and the remaining 24 indicator items scored a score of 1. Through the acquisition of these scores resulted in a total score of 491, so that it was included in the criteria for an appropriate instrument with minor improvements. For more details, the following table presents the results of the expert judgment evaluation instrument for the development of children aged 4-5 years.

Table 2. Results of Expert Judgment Results of Evaluation Instruments for Child Development Achievement Aged 4-5 years

| Core Competencies                          | Amount of Basic Competencies | Amount of Indicator Items | Validation Score | Expert Comments                                                                                                                                 |
|--------------------------------------------|------------------------------|----------------------------|------------------|---------------------------------------------------------------------------------------------------------------------------------|
| Spiritual Attitude Core Competencies (KI 1)| 2                            | 9                          | 27               | There are no special notes on all items                                                                                                   |
|                                            |                              |                            |                  | There are no special notes on 41 items; mention more specifically the targets of polite behavior; review several items that have the same meaning |
| Social Attitude Core Competencies (KI 2)   | 14                           | 53                         | 140              | There are no special notes on 83 items; specify the number of daily prayers memorized; review several items that have the same meaning |
|                                            |                              |                            |                  | replace some words with diction that are more appropriate in the context and meaning of the sentence; mention the target place more specifically according to the meaning of the sentence; ensure clear concept differences between clean behavior and healthy behavior; correct some meanings of sentences that are potentially biased and multiple interpretations |
| Knowledge Core Competencies (KI 3)        | 15                           | 87                         | 324              |                                                                                                                                              |
| Core Skills Competencies (KI 4)            | 15                           | 87                         | 324              |                                                                                                                                              |

Some of the themes of the comments found include: (1) specifically mentioning the number of daily prayer readings memorized; (2) adding targets to whom the polite behavior of children is shown; (3) re-examine several items that have the same meaning; (4) replace certain words with diction that is more appropriate in the context and meaning of the sentence; (5) mentioning the target location more specifically according to the meaning of the sentence; ensure clear concept differences between clean behavior and healthy behavior; correct some meanings of sentences that are potentially biased and multiple interpretations.
sentence; (6) ensuring clear concept differences between clean behavior and healthy behavior; (7) and several meanings of sentences that have the potential to be biased and multiple interpretations need to be improved. These comments form the basis for researchers in making improvements to indicator items according to the content and context of the sentence.

Specifically mentioning the number of the daily prayer reads memorized by children is a special note written by experts on one item of basic competence (KD) 3.1 & 4.1. Basic competence 3.1 reads to recognize daily worship activities, while basic competence 4.1 reads to perform daily worship activities. Thus children are not only expected to get to know but are also able to perform daily worship activities, in this case specifically on the ability to recite daily prayers.

Experts suggest specifically mentioning the number of daily prayers that need to be memorized by children aged 4-5 years—intended to be more operational when used to evaluate child development achievements on this item. However, this cannot be done because this indicator is made to be used universally in any religion. Therefore, this section does not make much improvement. The value of religion is important, because religion is proven to have a significant effect on various domains of early childhood development (Bartkowski et al., 2008). Religion and spirituality can be a protection for children (Bone & Fenton, 2015), religion and spirituality can be a protection for children (Mata-McMahon, 2019). Religion can also overcome stress and trauma in children from a bad event, such as child abuse (Brewer-Smyth & Koenig, 2014; Gall, 2006; Gall et al., 2007; Walker et al., 2009). Therefore, this section needs to be a learning goal in children aged 4-5 years.

Adding a target to whom the child's polite behavior is shown is a special note written by the expert on one item from basic competencies 3.2 & 4.2. Basic competence 3.2 reads recognizing good behavior as a reflection of noble morals, whereas KD 4.2 reads showing polite behavior as a reflection of noble morals. Manners' behavior is a reflection of a noble character. Children are not only expected to know but also show polite behavior to anyone as a reflection of noble morals. In one of the items in the KD pair, the expert suggested adding a target to whom the child could behave politely. Improvements were made to these indicator items so that specific behavioral targets were mentioned specifically, such as: to teachers, friends, and parents.

Courtesy behavior needs to be introduced and familiarized with the child since then, be it polite in speaking or polite in acting (Kusno, 2014; Ningsih et al., 2019). Early childhood awareness in courtesy behavior can be built from the family and the closest educational environment outside the family (Chen, 2017; Umaroh et al., 2017). The family will be the first environment of the child modeled on polite behavior. Then in the school environment, the teacher becomes a role model in behaving politely for children. Therefore, this is the expected outcome of children aged 4-5 years.

A special note from the next expert is to review several items that have the same meaning. These notes were found in several items from KD pairs that came from Core Competencies (KI) 3 and KI 4. According to experts, there are several meanings of vocabulary that are similar, so researchers are expected to pay close attention to ascertain the intent of these items. One example of the word in question is between "Accustomed to eating balanced nutritional food" in KD 2.1 with the item "Able to do activities that show healthy living behavior through exercise and eating balanced nutritional food" in KD 3.2-4.2.

After examining the phrases used in the indicators in KD 2.1 and KD 3.2-4.2, it can be explained that the emphasis on competence in KD 2.1 is a habit that has been consistently carried out by students in eating balanced nutritional food. While in KD 3.2 & 4.2 is a competency that emphasizes ability and skills. In this case, children may want and can eat foods with balanced nutrition and follow gymnastics as a form of exercise. However, if the
activity is not consistent within a certain period of time carried out by the child, then competence 2.1 cannot be said to be achieved.

Other items that are requested to pay close attention are because of the meaning of the vocabulary that needs to be ascertained in the difference in meaning. The items in question are items that contain the words healthy behavior and clean living behavior. In this section, researchers can ensure that the two phrases have different meanings but are related. Clean behavior supports healthy behavior.

Healthy behavior is behavior that is shown by doing enough physical activity, adequate rest and eating balanced nutritional food. Unlike the clean behavior, which is interpreted by the behavior shown in the activities of bathing, shampooing, cutting nails, washing hands properly, and maintaining environmental cleanliness (Julianti et al., 2018; Nurmahmudah et al., 2018). To achieve success in instilling clean behavior and healthy living behavior in children, influenced by parents (especially mothers), teachers, and friends (Berliana & Pradana, 2016; Kristanto et al., 2019; Rompas et al., 2018; Wulandari & Pertwi, 2018). Therefore, increasing PHBS behavior needs to be done from an early age (Aulina, 2018).

Subsequent comments on indicator items from KI 3 and KI 4 include: suggestions for replacing some words with diction that are more appropriate to the context and meaning of the sentence. It as well as some sentence meanings that are potentially biased and multiple interpretations that need to be improved, and then improvements have been made.

**CONCLUSION**

After going through the discussion above, it can be concluded that the development evaluation instruments for children aged 4-5 years that have been developed still need some improvement and after being repaired, can be used at the trial stage.

**ACKNOWLEDGMENT**

Thank you to Ahmad Dahlan University through the Institute for Research and Community Service (LPPM) for providing financial support to complete research on the development of this instrument.

**REFERENCES**

Aulina, C. N. (2018). Peningkatan Kesehatan Anak Usia Dini dengan Penerapan Perilaku Hidup Bersih dan sehat (PHBS) di TK Kecamatan Candi Sidoarjo. *AKSILOGIYA: Jurnal Pengabdian Kepada Masyarakat*, 3(1), 50. https://doi.org/10.30651/aks.v3i1.1480

Bagnato, S. J. (2007). *Authentic assessment for early childhood intervention: Best practices*. Guilford Press.

Bartkowski, J. P., Xu, X., & Levin, M. L. (2008). Religion and child development: Evidence from the Early Childhood Longitudinal Study. *Social Science Research*, 37(1), 18–36. https://doi.org/10.1016/j.ssresearch.2007.02.001

Berliana, N., & Pradana, E. (2016). Hubungan peran orangtua, pengaruh teman sebaya dengan perilaku hidup bersih dan sehat. *Jurnal Endurance*, 1(2), 75–80. https://doi.org/10.22216/jen.v1i2.984

Bone, J., & Fenton, A. (2015). Spirituality and child protection in early childhood education: A strengths approach. *International Journal of Children’s Spirituality*, 20(2), 86–99.

Brewer-Smyth, K., & Koenig, H. G. (2014). Could spirituality and religion promote stress resilience in survivors of childhood trauma? *Issues in Mental Health Nursing*, 35(4), 251–256. https://doi.org/10.3109/01612840.2013.873101
Chen, Y. (2017). Children's early awareness of the effect of interpersonal status on politeness. *Journal of Politeness Research, 13*. https://doi.org/10.1515/pr-2014-0016

Gall, T. L. (2006). Spirituality and coping with life stress among adult survivors of childhood sexual abuse. *Child Abuse & Neglect, 30*(7), 829-844. https://doi.org/10.1016/j.chiabu.2006.01.003

Gall, T. L., Basque, V., Damasceno-Scott, M., & Vardy, G. (2007). Spirituality and the current adjustment of adult survivors of childhood sexual abuse. *Journal for the Scientific Study of Religion, 46*(1), 101-117.

Gazali, A. (2012). Pengaruh Latar Belakang Pendidikan dan Pengalaman Mengajar Terhadap Profesionalisme Guru SMK Kompetensi Keahlian Teknik Audio-Video Se Kota Yogyakarta. *Yogyakarta: Universitas Negeri Yogyakarta.*

Gullo, D. F. (2005). *Understanding assessment and evaluation in early childhood education* (Vol. 95). Teachers College Press.

Hughes, R. T. (1996). Expert Judgment as an estimating method. *Information and Software Technology, 38*(2), 67–75. https://doi.org/10.1016/0950-5849(95)01045-9

Julianti, R., Nasirun, M., & Wembayarli, W. (2018). Pelaksanaan Perilaku Hidup Bersih dan Sehat (Phbs) di Lingkungan Sekolah. *Jurnal Ilmiah Potensia, 3*(2), 76–82.

Kristanto, H., Sucipto, S., & Atmojo, D. S. (2019). Hubungan pekerjaan ibu dengan kemampuan 8 indikator perilaku hidup bersih dan sehat pada siswa Madrasah Ibtidaiyah Kecamatan Majoroto Kota Kediri. *Prosidings Seminar Nasional 2018 “Peran Dan Tanggung Jawab Tenaga Kesehatan Dalam Mendukung Program Kesehatan Nasional”*, 144-148.

Kusno, A. (2014). Kesantunan Bertutur oleh Orang Tua kepada Anak di Lingkungan Rumah Tangga. *Dinamika Ilmu, 14*(1), 13–26.

Lad, B. K., & Kulkarni, M. S. (2010). A parameter estimation method for machine tool reliability analysis using expert Judgment. *International Journal of Data Analysis Techniques and Strategies, 2*(2), 155. https://doi.org/10.1504/IJDATS.2010.032455

Mata-McMahon, J. (2019). Finding connections between spirituality and play for early childhood education. *International Journal of Children’s Spirituality, 24*(1), 44–57. https://doi.org/10.1080/1364436x.2019.1619528

Neisworth, J. T., & Bagnato, S. J. (2004). The mismeasure of young children: The authentic assessment alternative. *Infants & Young Children, 17*(3), 198-212. https://doi.org/10.1097/0001163-200407000-00002

Ningsih, R., Boeriswati, E., & Muliaistuti, L. (2019). Language politeness: Pragmatic-sociocultural perspective. *Proceedings of the International Conference on Education, Language and Society, 426–431*. https://doi.org/10.5220/0008999904260431

Nitko, A. J. (1996). *Educational assessment of students*. ERIC.

Novianti, R., Puspitasari, E., & Chairilsyah, D. (2013). Pemetaan kemampuan guru PAUD dalam melaksanakan asesmen perkembangan anak usia dini di Kota Pekanbaru. *Jurnal SOROT, 8*(1), 95–104. https://doi.org/10.31258/sorot.8.1.2353

Nurmahmudah, E., Puspitasari, T., & Agustin, I. T. (2018). Perilaku hidup bersih dan sehat (PHBS) pada anak sekolah. *ABDIMAS: Jurnal Pengabdian Masyarakat, 1*(2), 46–52. https://doi.org/10.35568/abdimas.v1i2.327

Puckett, M. B., & Black, J. K. (2000). *Authentic assessment of the young child: Celebrating development and learning*. ERIC.

Pulkkinen, U., & Holmberg, J. (1997). A method for using expert Judgment in PSA (STUK-YTO-TR--129). Finnish Centre for Radiation and Nuclear Safety (STUK). http://inis.iaea.org/Search/search.aspx?orig_q=RN:28037507

Puspita, W. A., & Susilowati, S. (2015). *Pedoman Penilaian Hasil Pembelajaran*. P2TK PAUD Kementerian Pendidikan Nasional.
Expert Judgment of Learning Achievements Evaluation Instrument for Children Age 4-5 Years Old
DOI: 10.31004/obsesi.v5i1.524

Repulik Indonesia, K. P. dan K. N. (2014). *Peraturan Menteri Pendidikan dan Kebudayaan Nasional Nomor 146 Tahun 2014 tentang Kurikulum 2013 PAUD Lampiran V Pedoman Penilaian.*

Rohmadheny, P. S., Nuraini, F., & Setianingrum, I. (2019). *What are the Problems of Indonesian ECE Teachers in the Implementation of Learning Evaluation?* 24.

Rompas, R., Ismanto, A. Y., & Oroh, W. (2018). Hubungan peran orang tua dengan perilaku hidup bersih dan sehat anak usia sekolah di SD Inpres Talikuran Kecamatan Kawangkoan Utara. *Jurnal Keperawatan,* 6(1).

Sani, R. A. (2016). *Penilaian Autentik.* Bumi Aksara.

Suminah, E., Nugraha, A., Lestari, G. D., & Wahyuni, M. (2015). *Kurikulum Pendidikan Anak Usia Dini (Apa, Bagaimana, dan Mengapa?).* Direktorat Pembinaan Pendidikan Anak Usia Dini.

Umaroh, L., Kurniawati, N., & Sari, C. A. (2017). *An Investigation of Young Children's Politeness Principle.* *English Language and Literature International Conference (ELLiC) Proceedings,* 1(0), 315–319.

Walker, D. F., Reid, H. W., O'Neil, T., & Brown, L. (2009). Changes in personal religion/spirituality during and after childhood abuse: A review and synthesis. *Psychological Trauma: Theory, Research, Practice, and Policy,* 1(2), 130. https://doi.org/10.1037/a0016211

Werner, C., Bedford, T., Cooke, R. M., Hanea, A. M., & Morales-Nápoles, O. (2017). Expert Judgment for dependence in probabilistic modelling: A systematic literature review and future research directions. *European Journal of Operational Research,* 258(3), 801–819. https://doi.org/10.1016/j.ejor.2016.10.018

Wulandari, D. R., & Pertiwi, W. E. (2018). Pengetahuan dan peran orangtua terhadap perilaku hidup bersih dan sehat pada siswa SD di Kecamatan Kramatwatu Serang. *Jurnal Dunia Kesmas,* 7(4).