Journal of Innovation in Psychology, Education and Didactics
ISSN: 2247-4579, E-ISSN: 2392-7127
http://www.jiped.ub.ro/
Covered in : EBSCO, CEEOL, ProQuest, DOAJ, Scipio,
International Innovative Journal Impact Factor, CiteFactor,
EuroPub database, Open Academic Journals Index, ResearchBib,
Universal Impact Factor
2022, Volume 26, Issue 1, pages: 25-38, doi:10.29081/JIPED.2022.26.1.03

Quality Criteria Check –
Testing of a Scientific Giftedness Instrument for Preschool Age

Maria Sophie SCHÄFERS [1]*, Claas WEGNER [2]

Received: 24 January 2022/ Accepted: 11March 2022/ Published: 30 March 2022

Abstract

In the project 'Kleine BegInNa', a test instrument was developed based on the CHC theory to determine scientific giftedness in preschool. This could lead to individualized support adapted to differing abilities, which may have positive long-term effects on school performance. In this study, the quality (objectivity, reliability, and validity) of a natural scientific giftedness test was examined using a sample of 69 children aged between four and a half and six and a half years (m = 44.9 %). We found that our test meets the quality criteria for standardized test instruments in most analyses and can be used for future surveys.

Key words: Comparative study; Natural sciences; Natural scientific giftedness test; Preschool age; Quality check

How to cite: Schäfers, M.S., & Wegner, C. (2022). *Quality Criteria Check – Testing of a Scientific Giftedness Instrument for Preschool Age*. *Journal of Innovation in Psychology, Education and Didactics*. *Journal of Innovation in Psychology, Education and Didactics*, *26*(1), 25-38. doi:10.29081/JIPED.2022.26.1.03

---

[1] PhD Student Maria Sophie Schäfers, Bielefeld University, Universitätsstraße 25, 33615, Bielefeld, Germany, E-mail: maria_sophie.schaefers@uni-bielefeld.de

[2] PhD, Prof. Claas Wegner, Bielefeld University, Universitätsstraße 25, 33615, Bielefeld, Germany, E-mail: claas.wegner@uni-bielefeld.de

* Corresponding author

1. Introduction

The debate to start teaching science in kindergarten has long been seen as an issue within school education. Over the past few years there has been a growing demand to integrate science education into preschool (e.g., Lück, 2013) by having children explore scientific phenomena and problems. Pedagogical professionals in day-care centers should help children gain appropriate knowledge and aim to strengthen their *scientific literacy* (Bürgermeister et al., 2019; Leuchter, Saalbach & Hardy, 2014; Steffensky, 2017). This is possible given that children show a high level of interest and motivation to deal with scientific topics at an early age (Textor, 2012). In addition, their neuronal structures are interconnected in such a way that they can quickly link content together, which is important for learning science (Braun, 2012; Sodian & Thoermer, 2002; Schäfers & Wegner, 2021a). Current events such as climate change, the COVID-19 pandemic and rapid technological progress show the importance that children need to acquire scientific competency at an early stage to later deal with such problems (Kähler, Hahn & Köller, 2020; Schäfers & Wegner, 2022). Early support can also counteract the decline in interest in secondary school (Potvin & Hasni, 2014; Gebhard, Höttecke & Rehm, 2017) and the prevailing shortage of skilled workers (Schäfers & Wegner, 2020a). Therefore, day-care centers should take on a central role to offer individualized support and promote natural science education (Schäfers & Wegner, 2022). To ensure the success of this process, Leuchter and Saalbach (2014) have established characteristics which are needed to create a supportive learning environment:

- Gathering diverse experiences for the differentiation of children's pre-concepts of foundation of scientific knowledge (Schneider, Vamvakoussi & van Dooren, 2012);
- Creating cognitive conflicts for children to develop an openness to new ideas (Hardy et al., 2006);
- Behavioural adaptations of the pedagogical professional to include an active child-professional interaction (Bürgermeister et al., 2019; Gisbert, 2004);
- The use of explorative learning (Saalbach & Leuchter, 2014);
- Identifying prior knowledge to offer appropriate promotion (Carey, 2000).

Steffensky (2017) noted several studies that support the value of an early identification and the subsequent promotion of scientific abilities in kindergarten. A longitudinal study by Morgan and others (2016) demonstrated that general knowledge acquired at the end of kindergarten predicts academic achievement in the third grade of primary school. In addition, two studies from the Early Childhood Longitudinal Study (ECLS-K) indicate that basic mathematical knowledge as well as competencies in mathematics in every day kindergarten life can predict performance in science (Claessens & Engel, 2013; Saçkes, 2013). These results also hold true for the short term, as has been illustrated by Guo and others (2015), where they observed the direct influence of mathematical competency when learning basic biology. Regarding affective characteristics, studies have also demonstrated that support for interested and gifted children in science has a positive effect on long-term interest and scientific self-concept (Markowitz, 2004; Grosch, 2011; Hausamann, 2012; Schäfers & Wegner, 2020b). Therefore, there is an assumption that the earlier support begins, it will have stronger long-term effects will have on children and their behaviour.

2. Theoretical background

2.1. Conceptual framework of the project "Kleine BegInNa – small ones gifted in natural sciences"

The project 'Kleine BegInNa – small ones gifted in natural sciences' was established in 2019 and is a project of the Osthushenrich-center for gifted research in the biology faculty

(OZHB) at Bielefeld University (Wegner et al., 2020). The project's aim is to focus on the identification of scientific talent and to promote this in kindergarten (Schäfers & Wegner, 2022). It uses a holistic approach based on the principles of design-based research (Schäfers & Wegner, 2021b; Sandoval & Bell, 2004; Shavelson et al., 2003).

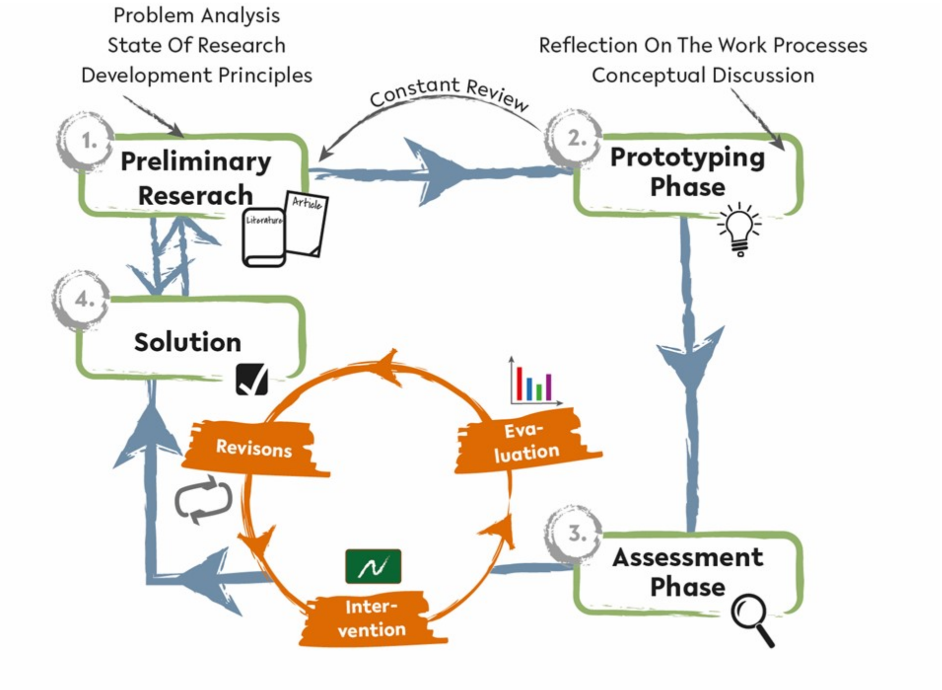

Figure 1. The design-based research approach in the project 'Kleine BegInNa'

This design-based research approach attempts to connect acquired scientific knowledge with practical problems by developing innovative solutions. Several researchers and stakeholders are involved in the research and development processes, which allows for a constant interaction between science and practice. The research design is divided into four main phases, though for the purpose of this study, only the first two are discussed in depth (see Figure 1). The preliminary phase consists of problem analysis and highlights previous findings within the current state of research. Although there are test instruments to assess scientific competence development in elementary school (e.g. the Science-P project, Hardy et al., 2010), an extensive literature review (Schäfers & Wegner, 2020a) covered two obstacles: first, there are hardly any tests to assess individual scientific abilities and second, that professionals in day-care centers are not sufficiently trained in the field of determining and promoting scientific giftedness (Fischnaller, 2012; Schuler, 2013; Bruns, 2014; Klemm et al., 2019). However, some studies have noted an influence of scientific promotion offered in day-care centers on later learning success in school (Claessens & Engel, 2013; Guo, Piasta & Bowles, 2015; Morgan et al., 2016).

2.2. Research concerns

From a prior literature review and previous known characteristics that create a conducive learning environment, the project 'Kleine BegInNa' has three main objectives (Schäfers & Wegner, 2022):

1. *Developing a diagnostic tool to determine scientific giftedness in kindergarten age*

By identifying abilities, individualized support can be offered to achieve long-term positive effects on scientific competence and school performance.

2. *Generating further training for pedagogical professionals with a focus on diagnostics and promotion in the natural sciences*

Target-oriented advanced training for pedagogical professionals can ensure the successful identification of scientific giftedness and any needed subsequent support.

3. *Implementation of promotion offers within kindergarten*

Depending on ability level, individualized interventions could be easily integrated into the daily routine in kindergarten, to continuously improve scientific competency (e. g. Schäfers & Wegner, 2020c; Schäfers & Wegner, 2021c; Schäfers & Wegner, 2021d).

To focus on the first goal, a natural scientific giftedness test (Schäfers & Wegner, 2022) and an accompanying observation sheet (Schäfers et al., 2020) were developed and examined. Therefore, this article focuses on the following questions:

Is our natural scientific giftedness test an objective, reliable, and valid test instrument to survey natural scientific abilities in preschool?

a. How is the objectivity of implementation, analysis and interpretation assessed?

b. Is the split-half method reliable?

c. Regarding the category system, which Cohen's kappa produces inter-rater reliability?

d. In terms of concurrent validity, how does the natural scientific giftedness test compare to existing instruments?

2.3. Development of the natural scientific giftedness test

The natural scientific giftedness test was developed based on the CHC theory of cognitive abilities (Carroll, 1993; Cattell, 1963; Horn, 1991; Horn & Blankson, 2005). The theory is a synthesis of theories from Cattell, Horn, and Carroll, and describes intelligence as a hierarchical system with three distinct levels (Flanagan & Dixon, 2013). The highest level consists of intelligence g, based on Spearman's g-factor (Spearman, 1904; see Figure 2). General abilities, such as fluid intelligence or auditory processing, are subordinate to the g-factor (see Figure 2). Each general ability is related to specific abilities that positively intercorrelate with each other. In total, there are 70 special abilities. For the giftedness test, general abilities relevant to science were filtered (marked in grey, orange and blue) based on the definition of natural scientific competencies by the IPN Kiel (IPN, n.d.), and subtests were developed for each of them.

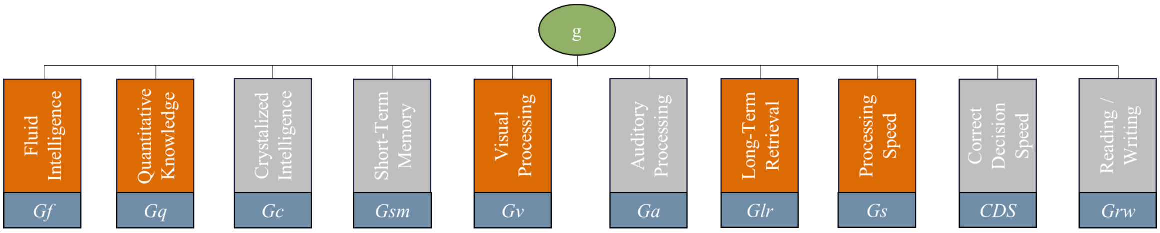

Figure 2. General intelligence and general abilities.

Subtests were partly developed on the basis of existing instruments for intelligence diagnostics. However, test contents were modified using a natural science focus. The following table (see Table 1) shows an overview of the natural scientific giftedness test subtests and which existing test instruments they are based on:

Table 1. Overview of the natural scientific giftedness test.

	Subtest	Content	Based on:
1	Long-term memory	The test leader starts reading a story about Iggi the hedgehog to the children. At the end of the test, the children must answer some questions about the story.	KABC-II (Kaufmann & Kaufmann, 2015)
2	Spatial-visual memory (farm)	The children get to know Max the dog, Mauzi the cat and Mimi the bird and must look for them on four different farms. The difficulty increases with each picture and the children have limited time.	IDS-P (Grob et al., 2013)
3	Number range	Small wooden cubes are used to test number skills. The children must give a certain number of cubes to the test leader.	IDS-P (Grob et al., 2013)
4	Fluid intelligence (flower series)	In this test, children must recognize the rules of different flower rows and complete them.	BIVA (Schaarschmidt et al., 2004)
5	Ordinality	Small wooden cubes are used to test ordinality skills. The children must give a certain cube to the test leader.	IDS-P (Grob et al., 2013)
6	Visual processing (animal cube)	With the large cube, the children should recognize the pairs of animals and deduce which animal is on the bottom of the cube.	HAWIVA-III (Ricken et al., 2007)
7	Quantity detection (animal herds)	To test their understanding of quantity, children familiarize themselves with different animals in the savannah. On individual pictures, they must identify where there are more individuals of a certain animal species.	IDS-P (Grob et al., 2013)
8	Fluid intelligence (Experiments)	The children must make assumptions about what happens during the experiments and why.	No model

3. Methodology

3.1. Setting

To establish the test instrument, a comparative study was conducted. After initial observations and shadowing day-care professionals to get to know the children and gain their trust, children completed a preliminary version of the natural scientific giftedness test (see Figure 3). Two weeks later, they were tested with comparable fragments of already validated aptitude tests. The main difference between the test fragments and our test is the focus on natural science. Results from both time points were compared to examine if the natural scientific giftedness test requires further modifications and whether the test meets the quality criteria.

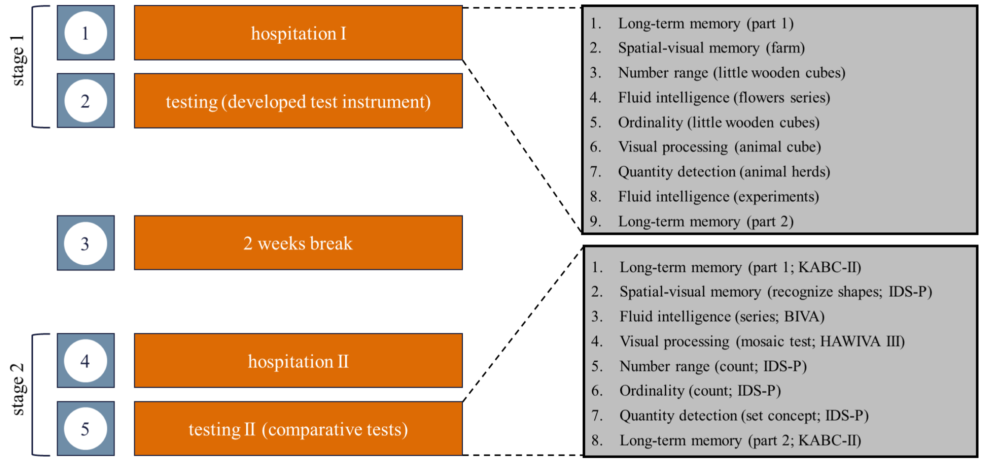

Figure 3. Study design.

3.2. Sample

The sample consisted of children (N = 69, male = 44.9%, average age = 5 years, 6 months) from three day care centers between 2020 and 2021.

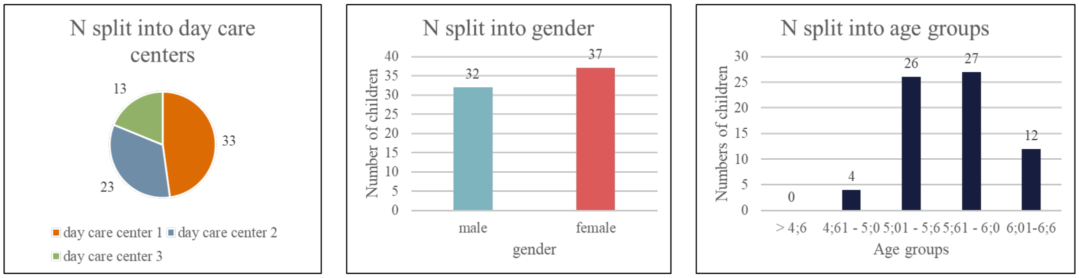

Figure 4. The sample of the study split into day care centers (a), gender (b) and age groups (c).

3.3. Statistical analysis

To examine the quality of the natural scientific giftedness test, we explored objectivity, reliability, and validity. To test reliability, split-half reliability and inter-rater reliability were conducted on the subtest 'Experiments'. Split-half reliability involves splitting the entire test into parts and correlating examinees' scores on both parts of the test instrument (Frey, 2018). To measure split-half reliability, separate items were combined into 'item twins' based on difficulty and selectivity. One item of each pair was randomly assigned to one half of the test and the second item to the other half. Inter-rater reliability (equivalence) is the extent to which the instrument produces consistent results across different users at a specified time (Gupta, 2012). Inter-rater reliability is indicated by Cohen`s kappa-coefficient, with values above 0.60 considered good and above 0.80 as nearly perfect (Grouven et al., 2007). Validity is checked by measuring concurrent validity. For this purpose, the results of the comparative study are tested for correlations. For the correlation analysis, the underlying data set contains the variables score t_0 (our developed test) and score t_1 (fragments of already validated tests) of the individual subtest pairings.

4. Results

The two halves of the test were examined for significant correlations as part of the split-half-reliability check (see Table 2). Results demonstrate that the comparison of all twin subtests, except for the pair *Animal herd I and II* (r = .20, p = .10, n = 69), produced significant correlations with moderately strong effects: *Story I and II* (r = .45, p < .001, n = 69), *Farm I and II* (r = .63, p < .001, n = 69), *Number + Ordinality abilities I and II* (r = .57, p < .001, n = 69), *Animal cube I and II* (r = .81, p < .001, n = 69), *Flower Series I and II* (r = .67, p < .001, n = 69) and *Experiments I and II* (r = .37, p < .01, n = 69). According to Cohen (1992), most of these are strong effects. According to Cohen (1992), most of these are strong effects.

Table 2. Correlations of split-half-reliability analysis.

		Story II	Farm II	N+O abilities II	Animal herds II	Animal cube II	Flower series II	Experi-ments II
Story I	Correlation Sig. (2-sided) N	.45*** *p* < .001 69						
Farm I	Correlation Sig. (2-sided) N		.63*** *p* < .001 69					
Number + Ordinality abilities I	Correlation Sig. (2-sided) N			.57*** *p* < .001 69				
Animal Herds I	Correlation Sig. (2-sided) N				.20 .10 69			
Animal Cube I	Correlation Sig. (2-sided) N					.81*** *p* < .001 69		
Flower Series I	Correlation Sig. (2-sided) N						.67*** *p* < .001 69	
Experiments I	Correlation Sig. (2-sided) N							.37** *p* < .01 69

Note: I = first half of natural scientific giftedness test; II = second half of natural scientific giftedness test. Pearson's correlation. * $p < .05$, ** $p < .01$, *** $p < .001$; N = sample size.

Since the subtest 'Experiments' required scores to be distributed, Cohen's kappa was determined to measure the internal consistency of the response category system. For this, two coders independently classified the children's answers into the category system and rated them with points from 0 to 2. Inter-rater reliability of 14 items in the subtest 'Experiments' was nearly perfect as indexed by Cohen's kappa (κ = .93). The percentage of agreement between the two independent coders was 96.38% (see Table 3).

Table 3. Inter-rater reliability.

		Coder 2		
		Category 2	Category 1	Category 0
Coder 1	Category 2	228	4	1
	Category 1	4	120	8
	Category 0	0	18	583

Note: Number of observed agreements: 931 (96.38% of the observations); Number of agreements expected by chance: 443.7 (45.93% of the observations)

Concurrent validity was assessed regarding validity; this is a form of criterion validity 'where there is evidence that scores from an instrument correspond to concurrently recorded external measures conceptually related to the measured construct' (Field, 2013, 872). Therefore, correlations between the self-developed subtests and the equivalent subtests of already validated measurement instruments were determined (see Table 4).

Table 4. Correlations for concurrent validity.

		Long-term memory Atlantis	Spatial-visual memory	Number range + Ordinality	Quantity detection	Mosaic test	Series
Story	Correlation	.37**					
	Sig. (2-sided)	< .01**					
	N	67					
Farm	Correlation		.44***				
	Sig. (2-sided)		< .001***				
	N		67				
Number + Ordinality abilities	Correlation			.44***			
	Sig. (2-sided)			< .001***			
	N			67			
Animal herds	Correlation				.04		
	Sig. (2-sided)				.74		
	N				67		
Animal cube	Correlation					.43***	
	Sig. (2-sided)					< .001***	
	N					67	
Flower series	Correlation						.40**
	Sig. (2-sided)						< .01**
	N						67

Note: Pearson's correlation. * $p < .05$, ** $p < .01$, *** $p < .001$; N = sample size.

Table 4 shows that the comparison of all paired subtests, except for the pair *Animal herd and Quantity detection* ($r = .04$, $p = .74$, $n = 67$), produced significant correlations with moderately strong effects: *Story and Long-term memory Atlantis* ($r = .37$, $p < .01$, $n = 67$), *Farm and Spatial visual memory* ($r = .44$, $p < .001$, n = 67), *Number + Ordinality abilities and Number range + Ordinality* ($r = .44$, $p < .001$, $n = 67$), *Animal cube and Mosaic Test* ($r = .43$, $p < .001$, $n = 67$), and *Flower Series and Series* ($r = .40$, $p < .01$, $n = 67$). Thus, the higher the score achieved on the developed subtest, the higher the score on the related subtest of the comparison test. According to Cohen (1992), these are moderately strong effects.

5. Discussion

We were interested in exploring the quality of our natural scientific giftedness test. Therefore, we calculated the reliability, validity, and objectivity of our test.

a. How is the objectivity of application, analysis and interpretation assessed?

A test is objective if it measures the characteristic to be measured independently of the test administrator, the test evaluator, and the interpretation of the results (Moosbrugger & Kelava, 2008). This can be divided into the objectivity of application, analysis, and interpretation (Lienert & Raatz, 1998). *The objectivity of application* of our test is guaranteed because it involves a standardized test procedure and is independent from the administration as precise instructions to use the test are given in a manual. These include, and are not limited to, time specifications, information about the material and linguistic impulses. It also specifies how to deal with questions to ensure uniform behaviour. Since children are not yet able to read and write, pen-and-paper along with computer-based tests are not an option for this age group. *The objectivity of analysis* is also verified, as a protocol is available for each test and does not depend on the test scorers. For many items, evaluation must differentiate between incorrect and correct, resulting in a high objectivity. Only the subtest 'Experiments' requires open answers to be evaluated. To maintain the quality of this test, a category system was developed to classify answers. This was checked as part of the reliability analyses.

b. Is the split-half method reliable?

A test is reliable if it accurately measures the characteristic without measurement errors (Moosbrugger & Kelava, 2008). First, the *reliability* of the entire test was examined using the *split-half method*. Our split-half method correlation analyses show significant relationships in all but one case with medium to strong effect sizes, indicating that the natural scientific giftedness test is a reliable instrument. The insignificant correlation for the subtest 'Animal herd' could be because the item twins differ too much in their selectivity and difficulty index. It must be noted that the split-half method does not ensure that the combined test halves are the same or that the test halves are the best possible similarity. Therefore, the split-half method assumes that the true reliability might be underestimated (Schermelleh-Engel & Werner, 2008). This means that the actual reliability of the natural scientific giftedness test must be estimated higher than the values indicate. Further surveys should check reliability again.

c. Regarding the category system, which Cohen's kappa produces inter-rater reliability?

Inter-rater reliability was found to be very high with a Cohen's kappa of $\kappa = .93$, indicating that the category system created was suitable and accurate. It also indicates a high degree of objectivity since it was created by different students within the project.

d. In terms of concurrent validity, how does the natural scientific giftedness test compare to existing instruments?

Validity is an integrated evaluative judgment about the extent to which the appropriateness and quality of interpretations and measures based on test scores, or other diagnostic procedures are supported by empirical evidence and theoretical arguments (Hartig, Frey & Jude, 2008). To test concurrent validity, a comparative study examined correlations between our natural scientific giftedness test to previously validated test instruments.

With one exception, we only observed significant correlations with a medium effect size. The lack of strong effects may be due to the fact that the tested ability domain was the same, but the context is different as we focused on the natural sciences. The comparison between the subtest ‘Animal herd’ and ‘Quantity detection’ was insignificant; this could be because the two subtests substantially differed in difficulty and complexity. In the validated test children had to decide which square contained the most objects. However, in our subtest they had to classify the animal species before they can count the similar animals.

6. Conclusion

This is the first study that focuses on developing a natural scientific giftedness test to determine individual abilities in preschool. Therefore, our results are limited when it comes to general statements about identifying natural scientific giftedness. This could be due to small sample sizes as the COVID-19 pandemic limited the extent to which researchers could implement surveys in kindergartens. Hopefully we will collect new data in the next months to further verify our test instrument. *Objectivity of interpretation* is met if the results can be interpreted in a standardized manner based on calibration samples or with a comparison to a reference sample. Because of the pandemic situation, the current sample is too small to construct standard tables, and therefore, further surveys must first be conducted. Nevertheless, this preliminary analysis shows that the test instrument is an objective, reliable and valid test instrument and meets the quality criteria. Future studies aim to conduct an exploratory factor analysis to explore the internal structure of the test instrument and determine correlations between the results of the first and second surveys (Schäfers et al., submitted).

Acknowledgements

The researchers would like to thank Simon Laumen for supporting the statistical operations and Arne Gerke for being our interrater. We also would thank Vivienne Litzke for the linguistic revision.

References

Braun, A. (2012). *Früh übt sich, wer ein Meister werden will - Neurobiologie des kindlichen Lernens. Weiterbildungsinitiative Frühpädagogische Fachkräfte*, WiFF Expertisen, Band 26. München.

Bruns, J. (2014). *Adaptive Förderung in der elementarpädagogischen Praxis: eine empirische Studie zum didaktischen Handeln von Erzieherinnen und Erziehern im Bereich Mathematik.* Münster u.a.: Waxmann.

Bürgermeister, A., Große, G., Leuchter, M., Studhalter, U., & Saalbach, H. (2019). Interaktion von pädagogischen Fachkräften und Kindern in naturwissenschaftlichen Lerngelegenheiten im Kindergarten. *Frühe Bildung*, 8 (1), 13-21.

Carey, S. (2000). Science education as conceptual change. *Journal of Applied Developmental Psychology*, 21, 13-19.

Carroll, J. (1993). *Human cognitive abilities: a survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Cattell, R. (1963). Theory of fluid and crystallized intelligence: a critical experiment. *Journal of Educational Psychology*, 54, 1-22.

Claessens, A., & Engels, M. (2013). How Important Is Where You Start? Early Mathematics Knowledge and Later School Success. *Teacher College Record*, 115 (6), 1-29.

Cohen, J. (1992). Statistical Power Analysis. *Current Directions in Psychological Science*, 1 (3), 98-101.

Fischnaller, S. (2012). *Zum Stand mathematischer Lernumgebungen in Kindergärten des Kindergartensprengels Brixen*. Laureatsarbeit: Bozen.

Flanagan, D., & Dixon, S. (2013). The Cattell-Horn-Carroll Theory of Cognitive Abilities. In C. Reynolds, K. Vannest, & E. Fletcher-Janzen (Eds.), *Encyclopedia of Special Education*. Hoboken: John Wiley & Sons Inc.

Field, A. (2013). *Discovering statistics using IBM SPSS Statistics*. Los Angeles/ London/ New Delhi/ Singapore/ Washington DC: SAGE Publishing Inc.

Frey, B. (2018). *The SAGE Encyclopedia of Educational Research, Measurement, and Evolution*. Los Angeles/ London/ New Delhi/ Singapore/ Washington DC: SAGE Publishing Inc.

Gebhard, U., Höttecke, D., & Rehm, M. (2017). *Pädagogik der Naturwissenschaften. Ein Studienbuch*. Wiesbaden: Springer Verlag.

Gisbert, K. (2004). *Lernen lernen. Lernmethodische Kompetenzen von Kindern in Tageseinrichtungen fördern*. Weinheim/ Basel: Beltz Verlag.

Grob, A., Reimann, G., Gut, J., & Frischknecht, M.-C. (2013). *Intelligence and Development Scales Preschool (IDS-P). Intelligenz- und Entwicklungsskalen für das Vorschulalter*. Bern: Huber.

Grosch, C. (2011). *Langfristige Wirkungen der Begabtenförderung*. Münster: LIT Verlag.

Grouven, U., Bender, R., Ziegler, A., & Lange, S. (2007). Der Kappa-Koeffizient. *Deutsches Medizinisches Rundschreiben*, 132, 65-68.

Guo, Y., Piasta, S., & Bowles, R. (2015). Exploring Preschool Children's Science Content Knowledge. *Early Education and Development*, 26(1), 125-146.

Gupta, A. (2012). *Assessmentinstrumente für alte Menschen. Pflege- und Versorgungsbedarf systematisch einschätzen*. Bern: Hans Huber Verlag.

Hardy, I., Jonen, A., Möller, K., & Stern, E. (2006). Effects of instructional support within constructivist learning enviroments for elementary school students´ understanding of „floating and sinking". *Journal of Educational Psychology*, 98(2), 307-326.

Hardy, I., Kleickmann, T., Koerber, S., Mayer, D., Möller, K., Pollmeier, J., Schwippert, K., & Sodian, B. (2010). Die Modellierung naturwissenschaftlicher Kompetenz im Grundschulalter. Projekt Science-P. In E. Klieme, D. Leutner, & M. Kenk (Eds.), *Kompetenzmodellierung. Eine aktuelle Zwischenbilanz des DFG-Schwerpunktprogramms*. (pp. 115-125). Weinheim/ Basel: Beltz.

Hartig, J., Frey, A., & Jude, N. (2008). Validität. In H. Moosbrugger, & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion*. (pp. 134-163). Heidelberg: Springer.

Hausamann, D. (2012). Extracurricular science labs for STEM talent support. *Roeper Review*, 34 (3), 170-182.

Horn, J. (1991). Measurement of intellectual capabilities: a review of theory. In K. McGrew, J. Werder, & R. Woodcock (Eds.), *Technical Manual. Woodcock-Johnson Psycho-Educational Battery-Revised*. (pp. 197-232). Chicago: Riverside.

Horn, J., & Blankson, A. (2005). Foundations for better understanding of cognitive abilities. In D. Flanagan, & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues*. (pp. 73-98). Guilford: The Guilford Press.

IPN (n.d.). *Bereich naturwissenschaftliche Kompetenz*. https://archiv.ipn.uni-kiel.de/PISA/index.html

Kähler, J., Hahn, I., & Köller, O. (2020). The development of early scientific literacy gaps in kindergarten children. *International Journal of Science Education*, 42 (12), 1988-2007.

Kaufman, A. S., & Kaufman N. L. (2015). *KABC-II: Kaufman Assessment Battery for Children-2*. Frankfurt: Pearson.

Klemm, J., Kohlhauf, L., Boone, W. Sodian, B., & Neuhaus, B. (2019). Förderung biologischer Beobachtungskompetenzen im Kindergarten. *Frühe Bildung*, 8 (1), 22-29.

Leuchter, M., & Saalbach, H. (2014). Verbale Unterstützungsmaßnahmen im Rahmen eines naturwissenschaftlichen Lernangebots in Kindergarten und Grundschule. *Unterrichtswissenschaft*, 42 (2), 117-131.

Leuchter, M., Saalbach, H., & Hardy, I. (2014). Designing science learning in the first years of schooling. An intervention study with sequenced learning material on the topic of ‘floating and sinking’. *International Journal of Science Education*, 36, 1751-1771.

Lienert, G., & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Beltz.

Lück, G. (2013). Förderung naturwissenschaftlicher Bildung. In M. Stamm, & D. Edelmann (Eds.), *Handbuch frühkindliche Bildungsforschung*. (pp. 557 – 573). Wiesbaden: Springer VS.

Markowitz, D. (2004). Evaluation of the Long-Term Impact of a University HighSchool Summer Science Program on Students` Interest and Perceived Abilities in Science. *Journal of Science Education and Technology*, 13 (3), 395-407.

Moosbrugger, H., & Kelava, A. (2008). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In H. Moosbrugger, & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion*. (pp. 7-26). Heidelberg: Springer.

Morgan, P., Farkas, G., Hillemeier, M., & Maczuga, S. (2016). Science Achievement Gaps Begin Very Early, Persist, and Are Largely Explained by Modifiable Factors. *Educational Researcher*, 45 (1), 18-35.

Potvin, P., & Hasni, A. (2014). Interest, motivation and attitudes towards science and technology at K-12-levels: a systematic review of 12 years of educational research. *Studies in Science Education*, 50 (1), 85-129.

Ricken, G., Fritz, A., Schuck, K.-D., & Preuß, U. (Eds.) (2007). *HAWIVA-III. Hannover-Wechsler-Intelligenztest für das Vorschulalter - III*. Bern: Huber.

Saçkes, M. (2013). Children’s Competencies in Process Skills in Kindergarten and Their Impact on Academic Achievement in Third Grade. *Early Education and Development*, 24 (5), 704-720.

Sandoval, W., & Bell, P. (2004). Design-Based Research Methods for Studying Learning in Context: Introduction. *Educational Psychologist*, 39 (4), 199-201.

Schaarschmidt, U., Ricken, G., Kieschke, U., & Preuß, U. (2004). *BIVA. Bildbasierter Intelligenztest für das Vorschulalter*. Göttingen: Hogrefe.

Schäfers, M. S., Höhne, M., Rehkemper, J., & Wegner, C. (2020). Darstellung eines Mixed-Methods-Ansatzes zur Bestimmung naturwissenschaftlicher Begabung im Kindergartenalter.

Diskurs Kindheits- und Jugendforschung / Discourse Journal of Childhood and Adolescence Research, 15 (4), 461-466.

Schäfers, M. S., Perleth, C., Bueno, C., Lübcke, H., & Wegner, C. (submitted). Wie lässt sich eine naturwissenschaftliche Begabung frühzeitig erkennen? Validierung eines naturwissenschaftlichen Begabungstests für den Elementarbereich.

Schäfers, M. S., & Wegner, C. (2020a). Diagnose und Förderung von naturwissenschaftlicher Begabung in der Kita. Darstellung des aktuellen Forschungsstands. *Diskurs Kindheits- und Jugendforschung / Discourse Journal of Childhood and Adolescence Research*, 15(1), 70-86.

Schäfers, M. S., & Wegner, C. (2020b). Long-term effects of the enrichment program "Kolumbus-Kids" - Differences in the career paths of former participants. *Journal of Gifted Education and Creativity*, 7(2), 73-82.

Schäfers, M. S., & Wegner, C. (2020c). Facettenreich. Die Augen der Fliege kennenlernen. *Entdeckungskiste*, 2020 (6), 58-59.

Schäfers, M. S., & Wegner, C. (2021a). Gehirn, Neuronen, Synapsen und Co. Eine neurobiologische Begründung für eine frühzeitige Förderung im Kindergartenalter. *frühe Kindheit*, 24 (4), 46-50.

Schäfers, M. S., & Wegner, C. (2021b). „Kleine BegInNa“: Ein Projekt zur Bestimmung und Förderung naturwissenschaftlicher Kompetenzen von Kitakindern. *ElFo - Elementarpädagogische Forschungsbeiträge*, 3(1), 9-20.

Schäfers, M. S., & Wegner, C. (2021c). Regenwurm-Alarm. *Draußenkinder*, 10 (2), 16-21.

Schäfers, M. S., & Wegner, C. (2021d). Motorikschleifen. Pädagogisches Material selbst hergestellt. *kindergarten heute*, 51 (6-7), 36-37.

Schäfers, M. S., & Wegner, C. (2022). Originalarbeit: Die naturwissenschaftlichen Fähigkeiten von Kindern entdecken. Vorstellung eines naturwissenschaftlichen Begabungstests für den Elementarbereich. *Frühförderung interdisziplinär*, 41 (1), 18-31.

Schermelleh-Engel, K., & Werner, C. (2008). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger, & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion*. (pp. 113-133). Heidelberg: Springer.

Schneider, M., Vamvakoussi, X., & van Dooren, W. (2012). Conceptual change. In N. Seel (Ed.), *Encyclopedia of the sciences of learning*. (pp. 735-738). New York: Springer.

Schuler, S. (2013). *Mathematische Bildung im Kindergarten in formal offenen Situationen. Eine Untersuchung am Beispiel von Spielen zum Erwerb des Zahlbegriffs*. Waxmann: Münster/ New York/ München/ Berlin.

Shavelson, R., Philipps, D. C., Towne, L., & Feuer, M. (2003). On the Science of Education Design Studies. *Educational Researcher*, 32 (1), 25-28.

Sodian, B., & Thoermer, C. (2002). Naturwissenschaftliches Denken im Grundschulalter: Die Koordination von Theorie und Evidenz. In K. Spreckelsen, K. Möller, & A. Hartinger (Eds.), *Ansätze und Methoden empirischer Forschung zum Sachunterricht*. (pp. 105-117). Bad Heilbrunn: Verlag Julius Klinkhardt.

Spearman, C. (1904). „General intelligence“, objectively determined and measured. *American Journal of Psychology*, 15, 201-293.

Steffensky, M. (2017). *Naturwissenschaftliche Bildung in Kindertageseinrichtungen. Weiterbildungsinitiative Frühpädagogische Fachkräfte*, WiFF Expertisen, Band 48. München.

Textor, M. (2012). *Bildung im Kindergarten*. Norderstedt: Books on Demand.

Wegner, C., Zehne, C., Wiese, I., & Schäfers, M. S. (2020). Introducing a Concept for Supporting Scientifically Gifted Students - the Osthushenrich-Center for Giftedness Research (OZHB) at the Department of Biology at Bielefeld University, Germany. *Journal of Gifted Education and Creativity*, 7 (1), 13-19.