Bridging Theory and Practice: Study of Informal Inferential Reasoning Instrument Development for Junior High School Students

Parhaini Andriani*[1]; Hesikumalasari[2]
[1,2]Universitas Islam Negeri Mataran
parhaini.andriani@uinmataram.ac.id[*1], hesikumalasari@uinmataram.ac.id[2]

Abstract

The goal of this study is to create a valid and reliable informal inferential reasoning tool. This is Research and Development (R & D) research that employs a modified Plomp model. An initial inquiry phase, a prototype phase, and an assessment phase comprise the development stages. The content validity test was conducted by two specialists, one in the field of Mathematics Education and one in the field of Mathematics. The trial featured a total of 28 class IX pupils from SMPN 8 Mataram. The product moment correlation was used to assess construct validity, while the Cronbach's Alpha coefficient was used to assess reliability. This study produced an informal inferential reasoning instrument with six question items based on three different contexts relevant to the challenge of comparing two groups of data. This instrument is suitable for usage, according to all validators. The trial findings also reveal that all items in this instrument meet the validity and reliability criteria.

Keywords: *Instrument Development; Informal Inferential Reasoning; Comparing two Groups of Data*.

Abstrak

Penelitian ini bertujuan untuk mengembangkan instrumen penalaran inferensial informal yang valid dan reliabel. Jenis penelitian ini adalah Reseach and Development (R & D) yang menggunakan model Plomp yang dimodifikasi. Tahapan pengembangan terdiri dari fase investigasi awal, fase prototipe, dan fase penilaian. Uji validitas konten melibatkan dua orang ahli yang terdiri dari 1 orang di bidang Pendidikan matematika dan 1 orang di bidang Matematika. Sebanyak 28 orang siswa kelas IX SMPN 8 Mataram dilibatkan dalam uji coba. Uji validitas konstruk diukur menggunakan korelasi product moment dan uji reliabilitas diukur menggunakan koefisien Alfa Cronbach. Hasil penelitian ini adalah instrumen penalaran inferensial informal yang mencakup 6 item pertanyaan dengan menggunakan 3 konteks yang berbeda yang berkaitan dengan masalah membandingkan dua kelompok data. Semua validator menyatakan bahwa instrumen ini layak digunakan. Hasil uji coba juga menunjukkan bahwa semua item dalam instrumen ini memenuhi kriteria valid dan reliabel.

Kata Kunci: Pengembangan Instrumen; Penalaran Inferensial Informal; Membandingkan Dua Kelompok Data.

*Correspondence:
Email: parhaini.andriani@uinmataram.ac.id

INTRODUCTION

The rise of the digital has brought about a lot of change in the world. With the help of technology, we are now able to do things that we are not able to do before. It has created a new opportunities for people to learn and grow. It is possible to utilize as much information from data as possible, due to the amount of data available even at your fingertips. As a result, more jobs and career advancement opportunities become available for data analysts. In order to be a qualified data analyst, a variety of qualifications are required. Hence, schools are required to prepare their students to be flexible thinkers, lifelong learners (Makar & Rubin, 2009) and capable of performing critical analysis of and evaluating data-driven claims (Ben-Zvi et al., 2015). When students are given time and guidance to explore and question their own hypotheses, their critical thinking and ability to make probabilistic generalizations from data significantly improve(Bråtalien & Naalsund, 2024). Therefore, schools and universities must cultivate students' critical and creative thinking skills in data comprehension. One of the most important things to develop in students is informal inferential reasoning.

A study comparing the statistics and probability curricula of Indonesia, Singapore, and the United States in 2013 found that the Indonesian curriculum placed greater emphasis on comprehending fundamental concepts and covered less material than those of the other two countries (Siregar, 2015). In Indonesia, statistical inference is not covered in the school curriculum. Several countries have introduced statistical inferences to primary school children, albeit informally. The reasoning of students in drawing statistical inferences informally is known as informal inferential reasoning.

Recent statistical education research has extensively explored informal inferential reasoning (Pratt & Ainley, 2008). This kind of reasoning is important for students at all levels because it helps them learn how to make predictions and choices based on real-world data (Garfield et al., 2015), and encouraging critical thinking(Bråtalien & Naalsund, 2024). This reasoning can improve students' understanding of formal statistical inference (Garfield et al., 2015; Jacob, 2013;

Schindler & Seidouvy, 2019; Zieffler et al., 2008). However, this reasoning has not been extensively studied in Indonesian schools or universities.

Inferential statistics, which is taught in universities, should start in school. Due to the importance of understanding inferential statistics (Zieffler et al., 2008) and the challenges of this type of reasoning (Dolor & Noll, 2015), it is widely accepted that the foundations of inferential statistics must be laid in early school years. In lower grades, students should generate samples (Doerr et al., 2017; Meletiou-Mavrotheris & Paparistodemou, 2015) and make statistical inferences using informal methods (Zieffler et al., 2008). There are two main ways to enhance inferential reasoning: firstly, by introducing it to younger students using an informal approach, and secondly, by developing formal inferential reasoning based on the informal methods(Lugo-Armenta & Pino-Fan, 2021). This research focuses on the first approach.

Andriani (2019) has developed a framework for assessing the informal inferential reasoning abilities of Junior High School pupils in the context of Indonesia. However, the development of school-useable instruments has not been thoroughly investigated. Even though the need for this instrument in the context of Indonesia is crucial for the development of students' statistical literacy. Therefore, this study will construct an informal inferential reasoning instrument based on Andriani's (2019) framework.

A framework for assessing junior high students' informal inferential reasoning skills has been developed by Andriani (2019). However, the development of school-appropriate instruments has not been thoroughly investigated as of yet. In the Indonesian context, this instrument is crucial for fostering the statistical literacy of students. Therefore, utilizing the framework established by Andriani (2019), this study will develop an informal inferential reasoning instrument.

RESEARCH METHODS

The type of research being conducted here is developmental research. The design that was used in this study is the development model that is outlined

by (Plomp, 2013). This model has three phases: the first phase is the initial investigation phase, the second phase is the prototype phase, and the third phase is the evaluation phase.

The trial phase subjects consisted a total of 28 students from ninth grade at SMPN 8 Mataram. The process of expert validation was conducted by one mathematics education expert and an one mathematics expert. The employed instrument is a format for expert validation that consists of expert evaluations of the validity of language and construct aspects in addition to qualitative evaluations relevant to the learning tools under development. Regarding the construction aspect, experts evaluate the quality of the problem, its alignment with the educational level, the sufficiency of information provided to solve the problem, and the context familiar to students. Concerning the language aspect, experts assess the adherence to proper language rules, communicativeness, and the avoidance of ambiguity. The assessments conducted by the experts provide a basis for the revision of the instrument before being tested.

RESULTS AND DISCUSSION

The purpose of this study is to develop a valid and reliable instrument for informal inferential reasoning that will be used with students in junior high school. An explanation of the findings of the research at each stage of development is provided below, along with an analysis of each of those particular findings.

1. Initial investigation phase

In this phase, a study is carried out regarding theories related to informal inferential reasoning. The results of the analysis in this phase obtained several relevant theories, theories about five levels of informal inferential reasoning, namely: (1) pre-structural; (2) pre-IIR; (3) naïve-IIR; (4) semi appropriate-IIR and (5) informal appropriate-IIR proposed by (Andriani, 2019) which is the result of the development of Goss' theory (2015). In addition, (Zieffler et al., 2008) proposed a theory regarding the tasks that can be employed to evaluate informal inferential reasoning: (1) predicting and

drawing population graphs; (2) comparing two samples of data; and (3) deliberating between two opposing models.

2. Prototype phase

In this phase, informal inferential reasoning task were developed. The form and format of the task chosen is an open ended question related to the task of comparing two groups of data. According to Table 1, the instrument includes 6 questions that use 3 types of data context. The instrument is also equipped with a scoring rubric which contains a scoring guide equipped with possible answers for each item. This rubric was developed based on the leveling developed by Andriani (2019). A summary of the question items developed is explained in Table 1. The results of development in this phase resulted in the Draft 1 instrument product.

Table 1. The Form of the Questions Developed in the Instrument

Item Number	Problem Context	Data Representation
1a and 1b	Compare the number of seeds in a packet of brand A and B sunflower seeds	Bar Chart
2a and 2b	Comparing the duration of gadget use by male and female students	Tale
3a and 3b	Compare student scores between those given worksheets and students who were not given worksheets	Dot plot

3. Evaluation phase

The informal inferential reasoning instrument is validated by two experts during this step. The two experts agreed that the final product was an usable instrument. Draft 2 of the instrument was created by incorporating validators' suggestions.

The development results were then tested at SMPN 8 Mataram using a Draft 2 instrument. We asked a total of 28 ninth graders to fill out this survey. The six items constructed for the instrument satisfy the valid criteria, as shown in Table 2, according to the validity test calculations conducted using SPSS in the form of product moment correlation values. The reliability test or internal consistency results, which were likewise computed using SPSS and presented

as Cronbach's Alpha coefficient values, reveal that all items achieved a value of $\alpha = 0.746$. Since the α value is more than or equal to 0.7, the constructed instrument can be considered to meet the dependability criterion.

Table 2. Instrument Validity Test Results

Item Number	r value	Validity Criteria
1a	0.508	Valid
1b	0.608	Valid
2a	0.716	Valid
2b	0.637	Valid
3a	0.752	Valid
3b	0.581	Valid
r table value = 0.3739		

The findings of the trial serve as a foundation for both the revision of the assessment rubric and the consideration of new questions for the instrument. This trial provided an empirical summary of potential student responses. Trials on instruments for informal inferential reasoning reveal that most students' inferential reasoning skills are lacking. This is shown in Table 3, which provides an overview of the findings from the informal inferential reasoning test. The majority of students are performing at the Pre-IIR level across all six questions, as seen in the table. Nonetheless, a number of pupils do well enough to be considered Semi Appropriate IIR. At this point, no one has attained the level of Appropriate IIR.

Table 3. Informal Inferential Reasoning Instrument Test Outcomes

Level	Item Number					
	1a	1b	2a	2b	3a	3b
Level 1 Pre-Structural	5	12	5	4	9	9
Level 2 Pre-IIR	19	11	20	17	6	10
Level 3 Naïve-IIR	3	5	3	7	11	6
Level 4 Semi-Appropriate IIR	1	0	0	0	2	3
Level 5 Appropriate IIR	0	0	0	0	0	0
Total	**28**	**28**	**28**	**28**	**28**	**28**

Based on the results of students' answers when testing the instrument, most of them only gave answers in the form of conclusions without writing

reasons or arguments that supported these conclusions. The language used to make generalizations or conclusions still uses deterministic sentences with a high level of certainty. Students who reach the semi-appropriate IIR level show reasoning that is aware of variability. There are students who are aware of outlier or outlier data which can be used as a basis for drawing conclusions. This development is analyzed based on the framework developed by Andriani (2019) which explains hierarchically the differences in reasoning in the aspects of understanding data, use of language in expressing uncertainty, understanding and use of context, and understanding variability, as explained in Table 4.

Table 4. Summary of Differences in Characteristics of Levels of Informal Inferential Reasoning (Adopted from Andriani, 2019)

Level	Understanding of data	The use of language in expressing uncertainty	Understanding and use of context	Understanding of variability
Level 1 Pre Structural	Local thinking by processing the information in the problem but the strategy is not relevant	Deterministic	-	-
Level 2 Pre-IIR	Local thinking by considering one aspect	Deterministic	Using data without context	-
Level 3 Naïve-IIR	Local thinking with some global thinking by considering several aspects	Limited use of probabilistic language	Using data more than context or vice versa, without having many implications for conclusions	Variability is poorly understood and ignored in conclusions.
Level 4 Semi Appropriate-IIR	Global thinking involves evaluating multiple factors and linking sample ideas to population characteristics.	Probabilistic language; near-exact conclusions	Context dominates data, or vice versa, with almost-correct conclusions.	Variability is understood and leveraged to make near-perfect generalizations.
Level 5 Appropriate-IIR	Global thinking by evaluating sample sizes, sampling procedures, and generating appropriate conclusions	Use of probabilistic language with appropriate conclusions	Used data and context properly with suitable conclusions	Appropriate understanding of variability and use in appropriate conclusions

CONCLUSION

An informal inferential reasoning instrument for students in junior high school was constructed as a result of this research. The instrument was developed through the following stages: (1) initial investigation phase; (2) prototype phase; and (3) evaluation phase. In order to compare two separate sets of data, this instrument consists of six different question items that are posed in three different contexts. This instrument is suitable for use, according to all of the validators." In addition, the findings of the study demonstrate that this instrument satisfies the requirements for validity and reliability. Teachers can use this instrument to investigate the inferential reasoning abilities of their students while they are in the classroom. Currently, the development of this instrument is restricted to a single category of tasks, specifically on comparing of two groups of data. Therefore, it is advised that future research be conducted in order to develop instruments similar to this one for various types of tasks, such as estimating and graphing populations and comparing two different models.

ACKNOWLEDGEMENT

This study is funded by the State Islamic University of Mataram for the 2022 fiscal year.

REFERENCES

Andriani, P. (2019). *Pelevelan Penalaran Inferensial Informal pada Siswa SMP/ MTs* [Disertasi]. Universitas Negeri Malang.

Andriani, P., Kurniawati, K. R. A., & Afriyani, D. (2022). A framework for Assessing Translation among Multiple Representations. *JTAM (Jurnal Teori Dan Aplikasi Matematika)*, *6*(2), 321. https://doi.org/10.31764/jtam.v6i2.7193

Ben-Zvi, D., Bakker, A., & Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics*, *88*(3), 291–303. https://doi.org/10.1007/s10649-015-9593-3

Bråtalien, M., & Naalsund, M. (2024). Informal Inferential Reasoning on Simple Linear Regression Encouraging Critical Thinking. *Nordic Journal of*

STEM Education, *8*(2), 73–87. https://doi.org/10.5324/njsteme.v8i2.4943

Doerr, H. M., Delmas, R., & Makar, K. (2017). A Modelling Approach To The Development Of Students' Informal Inferential Reasoning. *Statistics Education Research Journal*, *16*(2), 86–115. https://doi.org/10.52041/serj.v16i2.186

Dolor, J., & Noll, J. (2015). Using Guided Reinvention To Develop Teachers' Understanding Of Hypothesis Testing Concepts. *Statistics Education Research Journal*, *14*(1), 60–89. https://doi.org/10.52041/serj.v14i1.269

Garfield, J., Le, L., Zieffler, A., & Ben-Zvi, D. (2015). Developing students' reasoning about samples and sampling variability as a path to expert statistical thinking. *Educational Studies in Mathematics*. https://doi.org/10.1007/s10649-014-9541-7

Jacob, B. L. (2013). *The Development of Introductory Statistics Students' Informal Inferential Reasoning and Its Relationship to Formal Inferential Reasoning* [Doctoral Disertation]. Syracuse University.

Lugo-Armenta, J. G., & Pino-Fan, L. R. (2021). Inferential Reasoning of Secondary School Mathematics Teachers on the Chi-Square Statistic. *Mathematics*, *9*(19), 2416. https://doi.org/10.3390/math9192416

Makar, K., & Rubin, A. (2009). A Framework for Thinking about Informal Statistical Inference. *Statistics Education Research Journal*, *8*(1), 82–105.

Meletiou-Mavrotheris, M., & Paparistodemou, E. (2015). Developing students' reasoning about samples and sampling in the context of informal inferences. *Educational Studies in …*. https://doi.org/10.1007/s10649-014-9551-5

Plomp, T. (2013). Educational Design Research: An Introduction . In T. Plomp & N. Nieven (Eds.), *An Introduction to Educational Design Research* (pp. 11–50). Netherlands Institute for Curriculum Development (SLO).

Pratt, D., & Ainley, J. (2008). Introduction to Special Issue on Informal inferential reasoning. *Statistics Education Research Journal*, *7*(2), 3–4.

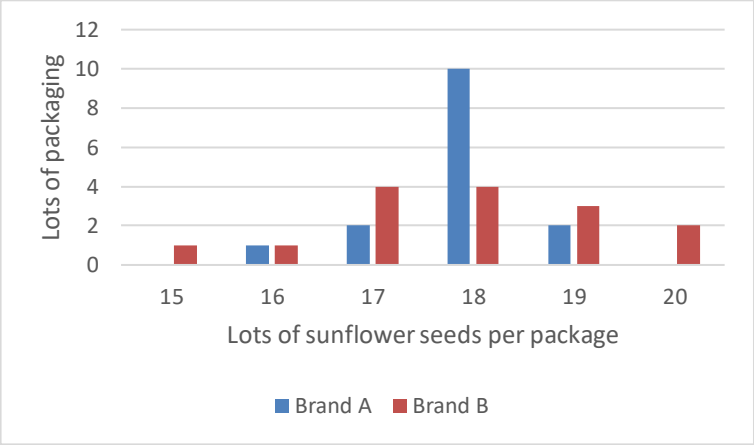
Schindler, M., & Seidouvy, A. (2019). Informal inferential reasoning and the social: understanding students' informal inferences through an inferentialist epistemology. *Topics and Trends in Current Statistics Education …*. https://doi.org/10.1007/978-3-030-03472-6_7

Siregar, E. M. R. S. (2015). A Comparative Study of International Mathematics Curriculum Materials and Standard: Indonesia, Singapore and The US Common Core State Standards. *Research and Creative Activities Poster Day*, 153.

Zieffler, A., Garfield, J., Delmas R., & Reading, C. (2008). A Framework To Support Research On Informal Inferential Reasoning. *Statistics Education Research Journal*, *7*(2), 40–58.

Appendix
Informal Inferential Reasoning Task

1. A total of 15 sunflower seed packages, labeled A and B, weighed 13 grams each. The contents of each package were examined. The data is represented in the below graph.



   a. If your friend is a fan of sunflower seeds, which brand of sunflower seeds would you recommend to him/ her? Explain your reasons.
   b. If more samples were taken from each of brand A and B, do you think there would be a difference in the overall quality of brands A and B? Explain your reason

2. A teacher conducted a survey about the daily use of gadgets among his 15 year old students. The following is data about the duration of gadget use from 24 male students and 21 female students.

Table 1. Duration of gadget use (samples = 45 students)

Duration (in minute)	Men	Women
1 – 60	0	1
61 – 120	0	1
121 – 180	5	7
181 – 240	11	10
241 – 300	7	2
More than 300	1	0
Total	**24**	**21**

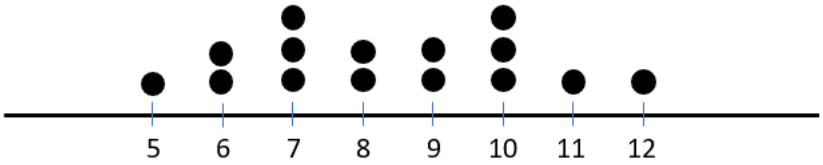
   a. What conclusions can be drawn from Table 1 regarding the gadget usage of male and female students? Explain your reasons.

Table 2. Duration of gadget use (samples = 81 students)

Duration (in minute)	Men	Women
1 – 60	1	1
61 – 120	4	7
121 – 180	9	12
181 – 240	11	15
241 – 300	9	9
More than 300	2	1
Total	**36**	**45**

b. If there is an additional sample as in Table 2, what conclusions do you think can be drawn regarding gadget use between male and female students? Explain your reasons.

3. A community health center officer measures the weight of a toddler who came to check at the beginning of the month at the community health center. The following is data on the weight of 15 toddlers who were examined from Monday to Friday.



Toddler's weight in kg

a. If on Saturday there are additional weight measurements of 8 toddlers, add the possibility of 10 pieces of data to the graph above! Explain your reasons for choosing this data.
b. What do you think the possible shape of a dot plot graph for checking a toddler's weight at the end of the month would be? Draw the possible shapes of the dot plot graph and explain your reason.