Effect of Reliability and Validity Testing in Guidance and Counseling Research

Tiara Anggun Mauladi[1],Ahman[2], Recita Widyan Morteza[3], Febi Dwi Fahriza[4], Samsul Aripin[4]*

Universitas Pendidikan Indonesia, Indonesia
*Correspondence: E-mail: angguntiara21@upi.edu

A B S T R A C T

Reliability and validity testing are essential steps in guidance and counseling research to ensure the quality of research instruments. This study aims to analyze the effect of reliability and validity testing on the quality of research results in the field of guidance and counseling. The method used is a systematic literature review (SLR) by analyzing 10 articles published in the period 2020-2024. The results show that valid and reliable instruments contribute significantly to the accuracy, consistency, and relevance of research data. Validity ensures that the instrument measures the intended construct, while reliability guarantees the consistency of measurement results. Instruments that do not meet these criteria can lead to data bias, unreliable conclusions, and inappropriate interventions. Therefore, systematic development instruments, involving experts, and considering local contexts are important factors in improving research quality. This study emphasizes the importance of validity and reliability in measurement to support decision making based on accurate data in the field of guidance and counseling.

A R T I C L E   I N F O

1. INTRODUCTION

Research is very important for guidance and counseling teachers because it allows them to gain a broader perspective, increase accountability, develop new ideas and approaches, support the application of guidance and counseling in new areas such as organizations, health care, and response to trauma, and contribute to the personal and professional development of counselors (McLeod, 2003). Reliability and validity testing is very important in guidance and counseling research, to determine the quality of the instrument on the accuracy, relevance, and representativeness of the data generated, and impact on the reliability and validity of the research results (Barahona et al., 2023). Machfoedz (2009) suggests that validity is accuracy and accuracy, in the world of research is valid or valid. Then, according to Sürücü and Maslakçi (2020) reliability refers to the stability of the measuring instrument used and consistency over time. Therefore, validity and reliability tests are important stages in guidance and counseling research to ensure the instrument is able to provide consistent results in measurement.

Quantitative research is an investigation of a social problem that is based on testing a theory involving variables, measured with numbers, and analyzed using statistical procedures to determine whether the predictive generalizations of the theory are true (Houser, 2020). Descriptive statistical methods are used to describe data characteristics, while inferential statistics are used to test relationships or differences between variables, such as t-tests, regression analysis, and analysis of variance (Creswell, 2012). Research with statistical-based quantitative methods is a research approach that focuses on collecting and analyzing numerical data to answer research questions. The data obtained is analyzed using statistical techniques to find specific patterns, relationships, or effects in the research context.

The importance of valid and reliable instruments in guidance and counseling research is because, accurate measurement plays a key role. Errors in measurement can lead to inappropriate conclusions, impact the quality of services, and even risk the well-being of clients. For example, an invalid instrument may produce biased data, resulting in interventions that are not appropriate for the client's condition (Creswell, 2012). In addition, unreliable instruments can produce inconsistent data, making it difficult to replicate or validate research results (Ary et al., 2010).

Many studies in the field of counseling use instruments developed without an adequate testing process. This is often due to limited resources or knowledge about instrument development. Research conducted by Kerlinger and Lee (2000) states that many of the instruments used have not gone through an adequate validity and reliability testing process, thus affecting the quality of the research itself. For example, school counselors often rely on commonly available measurement tools without checking their suitability for the local context (Suryani, 2021). The impact on research quality of lack of validity and reliability of instruments can lead to: (1) data errors, the data collected does not reflect the actual conditions; (2) inference bias, the research results become unreliable, so that decisions made based on the data are inappropriate (Cohen et al., 2007); (3) harm to clients, interventions based on invalid data can worsen the client's condition (Ary et al., 2010).

The focus of this study is to analyze the influence of reliability and validity testing in ensuring the quality of research results in the field of guidance and counseling. Validity and reliability testing directly affect the quality of data and interpretation of research results. If validity and reliability are not tested, the research results only reflect a small part of the construct being measured, resulting in biased and inaccurate data (Cohen et al., 2007). Valid and reliable instruments produce data that are: (1) Relevant, i.e. the data reflects the reality

of the research subject; and (2) Reliable, i.e. the research findings can be used to make appropriate decisions in counseling services (Azwar, 2016).

In this context, it is important to understand various instrument development models that have proven effective in improving the validity and reliability of measuring instruments used in education. One model that is often used is the 4D model, which consists of the stages of define, design, develop and disseminate. This model has been applied in various studies, such as in the development of humanistic literacy instruments and critical thinking skills (Suradipa, 2022). Research focusing on the development of Higher Order Thinking Skills (HOTS) assessment instruments also provides important insights. For example, research conducted by Budiman and Jailani (2014) and research conducted by Wulandari and Radia (2021) showed that HOTS instruments developed must consider the context and characteristics of students. This shows that instrument development must not only be valid and reliable, but also relevant to the needs and conditions of students in the field. In order to improve the quality of research instruments, it is also important to involve experts in the validation process. Research by Laksono (2020) emphasizes the importance of logical and empirical validity in instrument development, where input from experts can help ensure that the instrument developed can actually measure what it is intended to measure.

2. METHODS

This research uses the systematic literature review (SLR) method. Systematic review is a research method that collects primary research results to convey more complete and balanced facts (Hadi et al., 2018). The purpose of SLR is to identify, review, evaluate, and understand in depth all available research related to a particular topic or phenomenon (Triandini et al., 2019).Data collection carried out in this study is to search for literature from the Publish Or Perish (POP) application with the google scholars database, as for the limitations of the criteria or keywords used, namely (Reliability * and Validity * Guidance and Counseling Research). The articles found were 10 articles and those collected covered the time span of the last 5 years (2020-2024). The articles were then grouped based on predetermined criteria. The analysis in this study uses synthesis (summarizing) the findings in scientific articles with reference to the predetermined inclusion criteria as shown in **Table 1.**, namely:

Table 1. Inclusion criteria.

Research articles that discuss the concepts of reliability and validity in the context of guidance and counseling
Articles published in the last 5 years (2020-2024)
National Journal
Studies that explicitly explain the methodology of reliability and validity testing and its impact on research results

3. FINDINGS AND DISCUSSIONS
3.1 Findings

The results of this study present some literature related to the effect of testing the reliability and validity of an instrument in guidance and counseling research, the research journals obtained, which can be seen in the **Table 2.** below:

Table 2. Research articles that meet the inclusion criteria.

No	Article Title and Author	Description
1.	"Reliability and Validity Test of Depression Anxiety Stress Scales 21 (DASS-21) on College Students" by (Arjanto, 2021)	The research article tested the Depression Anxiety Stress Scales 21 (DASS-21), where reliability and validity testing carried out thoroughly can improve the quality of research instruments, so that research results become more accurate and reliable.
2.	"Validity and Reliability of Self-Adjustment Questionnaire for High School Students" by (Tussaadah et al., 2021)	The research article tested the self-adjustment questionnaire of high school students, where the reliability was 0.718 which was categorized as high. The high reliability of the research instrument contributes significantly to the consistency of the data collected, thus supporting the generalizability of the research findings.
3.	"Religious Moderation Scale (RM-S): Validity, Reliability and Implications in Guidance and Counseling" by (Monalisa et al., 2023)	The research article tests the Religious Moderation Scale (RM-S), where reliability and validity testing carried out thoroughly can improve the quality of research instruments, so that research results become more accurate and reliable.
4.	"Validity and Reliability of Emotional Maturity Questionnaire" by (Ramdani et al., 2023)	The research article tests the emotional maturity questionnaire, where the reliability score is 0.882 in the high category. so that high reliability on research instruments contributes significantly to the consistency of the data collected, thus supporting the generalization of research findings.
5.	"Validity and Reliability Test of Social Interaction Questionnaire" by (Taopik et al., 2023)	The research article tests the social interaction questionnaire, where the reliability test gets a value of 0.743, which means it is included in the high category. so that high reliability on research instruments contributes significantly to the consistency of the data collected, thus supporting the generalization of research findings.

6.	"Development of Inventory of Aggressive Behavior Tendencies of Junior High School Students" by (Rozzaqyah et al., 2021)	The research article develops an inventory of the tendency of aggressive behavior of junior high school students, where the results of inventory development show the Cronbach alpha reliability value of 0.858 with high reliability criteria and 34 inventory items are said to be reliable with the value of r count> r table, which is> 0.138. Therefore, it can be concluded that the inventory of aggressive behavior tendencies produced is valid and reliable so that it is expected to be used by schools and school counselors.
7.	"Development of Self Awareness Character Measurement Instrument in Pre Adolescence" by (Nurislami et al., 2021)	The research article develops a character measurement instrument for self awareness in pre-adolescence, where the Croanbach Alpha value is 0.933. The results of the reliability calculation are then compared with the value of the r table with the number N = 64 (number of respondents) at the 5% significance level, which is 0.244 so that the results show 0.933> 0.244. Thus, the instrument for measuring the character of self awareness is said to be reliable and has a very reliable classification as a measuring instrument for the character of self awareness of pre-adolescent students in junior high school.
8.	"Validity and Reliability Test of Learning Independence Questionnaire" by (Isma et al., 2023)	The research article tests the learning independence questionnaire, where the reliability test gets a value of 0.908, which means that he is in the highest category. so that high reliability on research instruments contributes significantly to the consistency of the data collected, thus supporting the generalization of research findings.
9.	"Validity and Reliability of the Psychological Self-Confidence Scale to Measure the Self-Confidence Level of Kediri City Vocational School Students" by (Krisphianti et al., 2020)	The research article tests the questionnaire Student self-confidence level, where the reliable instrument is 0.814. Based on this test, the psychological scale of self-confidence can be used as a measuring tool for the self-confidence of students of SMK Kediri City in the next stage of research.
10.	"Validity and Reliability of Academic Procrastination Questionnaire" by (Nuryana et al., 2022)	The research article tests the academic procrastination questionnaire, where the reliability value is at 0.911 with a very high category, which means that the procrastination questionnaire can be used anytime and anywhere to capture information about academic procrastination.

3.2 Discussions

High reliability in research is a key factor that positively impacts the findings. In the context of research, reliability refers to the consistency and stability of results obtained from

measurement instruments. When an instrument has high reliability, the researcher can be confident that the results obtained are not coincidental, but rather reflect the true phenomenon. Leppink and Pérez-Fuster (2017) emphasize the importance of test-retest reliability in replication research, where variations in results caused by unreliable instruments can result in wasted replication efforts. This suggests that high reliability not only increases confidence in the research results, but also in the validity of the findings.

In addition, research by Zhang et al. (2024) showed that good test-retest reliability, as found in the Chinese language version of the Community's Self-Efficacy Scale, provided a satisfactory kappa coefficient (0.754) in two measurements taken within a two-week time span. This suggests that high reliability can help in distinguishing within-sample results, which is particularly important in the context of research involving specific populations. Thus, high reliability contributes to the accuracy and reliability of research results, which in turn strengthens the arguments and conclusions drawn.

Furthermore, Byers-Heinlein et al (2022) highlighted that to improve effect size and measurement reliability in infant research, it is important to consider measurement validity and ecology. This suggests that high reliability should be combined with validity to produce more robust and reliable findings. Kucker and Chmielewski (2022) also emphasize that considering different types of reliability and validity indices will result in better measurement tools for researchers. Thus, high reliability not only has an impact on research results, but also on the quality of the measurement tools used. In the context of qualitative research, William (2024) underlined the importance of reliability and validity in determining the value of scientific research. In research involving qualitative data, the reliability and validity of the results are crucial to ensure that the findings are trustworthy and can be applied to a broader context. This is in line with Iksan and Saper (2016) view that although it is difficult to achieve validity and reliability in qualitative research approaches, it can be achieved through organized and ethical data management. Therefore, high reliability not only increases confidence in the research results, but also strengthens the credibility of the research as a whole.

The characteristics of a good research instrument can be seen from the results of the validity and reliability tests conducted. Validity refers to the extent to which the instrument measures what it is supposed to measure, while reliability relates to the consistency of the measurement results. A good instrument must have both of these aspects to ensure that the data obtained is reliable and accurate.

First, instrument validity can be tested through several methods, including content validity and construct validity. Mudanta et al. (2020) explained that a good assessment instrument must go through a validity analysis to ensure that each item in the instrument is relevant and in accordance with the research objectives. In addition, Amanda et al. (2019) emphasized the importance of the validity test to ensure that the questionnaire used actually measures the level of people's political participation. High validity results indicate that the instrument is reliable in collecting relevant data.

Second, instrument reliability is also an important indicator of the quality of research instruments. In a study by Novitasari et al. (2017), the clinical competence self-assessment instrument of medical students was tested for reliability and showed satisfactory results, indicating that the instrument can be relied upon to measure the competence in question. High reliability test results, as found by Wulandari and Radia (2021), where the responsibility attitude assessment instrument showed a reliability value of $\alpha = 0.944$, indicate that the instrument is very reliable.

Furthermore, a good instrument must also consider the level of difficulty and distinguishing power of each item. Item analysis showed that a good instrument should have a variety in difficulty levels and sufficient differentiating power to distinguish between respondents of different abilities Zikra and Laras (2021). This is important to ensure that the instrument can provide useful and relevant information.

Finally, good instrument development should involve a systematic and planned process, so instrument development should follow a clear methodology, including validity and reliability testing before the instrument is used in research. This process ensures that the resulting instrument is not only valid and reliable, but also appropriate to the context and purpose of the research.

4. CONCLUSION AND RECOMMENDATION

Thus, judging from the ten literatures, it can be concluded that high reliability has a significant impact on the research findings. It not only increases confidence in the results obtained, but also contributes to the overall validity and credibility of the research. Researchers should continuously strive to improve the reliability of the measurement tools used in their research to ensure that the resulting findings are reliable and applicable in a broader context. The characteristics of a good research instrument can be identified through the results of validity and reliability tests which show that the instrument is reliable, relevant, and able to accurately measure the intended variables. Researchers should always ensure that the instruments used meet these criteria to produce valid and reliable findings.

5. ACKNOWLEDGEMENTS

6. REFERENCES

Amanda, L., Yanuar, F., and Devianto, D. (2019). Uji Validitas dan reliabilitas tingkat partisipasi politik masyarakat Kota Padang. *Jurnal Matematika UNAND*, *8*(1), 179.

Andrés Cisneros Barahona, Luis Márques Molías, NIcolay Samaniego Erazo, Catalina Mejía Granizo, and Gabriela de la Cruz Fernández. (2023). *Data availability. multivariate data analysis. validation of an instrument for the evaluation of teaching digital competence.* [Dataset]. Zenodo. https://doi.org/10.5281/ZENODO.10055380

Arjanto, P. (2021). Uji Reliabilitas dan validitas Depression Anxiety Stress Scales 21 (DASS-21) pada mahasiswa. *Jurnal Psikologi Perseptual*, *7*.

Ary, D., Jacobs, L. C., Razavieh, A., and Ary, D. (2010). *Introduction to research in education* (8th ed). Wadsworth.

Azwar, S. (2016). *Reliabilitas dan validitas instrumen psikologi*. Pustaka Pelajar.

Budiman, A., and Jailani, J. (2014). Pengembangan instrumen asesmen higher order thinking skill (HOTS) PAda mata pelajaran matematikA SMP Kelas VIII Semester 1. *Jurnal Riset Pendidikan Matematika*, *1*(2), Article 2. https://doi.org/10.21831/jrpm.v1i2.2671

Byers-Heinlein, K., Bergmann, C., and Savalei, V. (2022). Six solutions for more reliable infant research. *Infant and Child Development*, *31*(5), e2296.

https://doi.org/10.1002/icd.2296

Cohen, R. J., Schneider, W. J., and Tobin, R. (2007). Psychological Testing and assessment. *McGrawHill*.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed). Pearson.

H. Iksan, Z., and Saper, M. N. (2016). The Process of validity and reliability for verbal questioning research in the classroom. *Global Journal of Educational Studies*, *2*(1), 14. https://doi.org/10.5296/gjes.v2i1.8755

Hadi, P., Sinring, A., and Aryani, F. (2018). Pengaruh permainan tradisional dalam meningkatkan keterampilan sosial siswa SMP. *Jurnal Kajian Psikologi Pendidikan Dan Bimbingan Konseling*, *4*(1).

Houser, R. A. (2020). *Counseling and educational research fourth edition* (4th ed.).

Isma, S., Sobari, T., and Yuliani, W. (2023). Uji validitas dan reliabilitas angket kemandirian belajaR. *FOKUS : Kajian Bimbingan dan Konseling Dalam Pendidikan*, *6*(6), 509–516. https://doi.org/10.22460/fokus.v6i6.11561

Kerlinger, F. N., and Lee, H. B. (2000). *Foundations of behavioral research*. Harcourt College Publishers.

Kouam Arthur William, F. (2024). Mastering validity and reliability in academic research: meaning and significance. *International Journal of Research Publications*, *144*(1). https://doi.org/10.47119/IJRP1001441320246160

Krisphianti, Y. D., Nora Yuniar Setyaputri, and Galang Surya Gumilang. (2020). Validitas dan reliabilitas skala psikologis percaya diri untuk mengukur tingkat percaya diri siswa SMK Kota Kediri. *PINUS: Jurnal Penelitian Inovasi Pembelajaran*, *6*(1), 57–65. https://doi.org/10.29407/pn.v6i1.14551

Kucker, S. C., and Chmielewski, M. (2022). Yes, but … We need reliability to advance infant work, but there's more to consider too: Important nuances of reliability and the need to include validity. *Infant and Child Development*, *31*(5), e2325. https://doi.org/10.1002/icd.2325

Laksono, P. J. (2020). Pengembangan three tier multiple choice test pada materi kesetimbangan kimia mata kuliah kimia dasar lanjut. *Orbital: Jurnal Pendidikan Kimia*, *4*(1), 44–63. https://doi.org/10.19109/ojpk.v4i1.5649

Leppink, J., and Pérez-Fuster, P. (2017). We need more replication research – Aandnbsp;case for test-retest reliability. *Perspectives on Medical Education*, *6*(3), 158–164. https://doi.org/10.1007/S40037-017-0347-Z

Machfoedz, I. (2009). *Metodologi penelitian (5th ed.).* Fitramaya.

McLeod, J. (2003). *Doing counselling research* (2nd ed). Sage Publications.

Monalisa, M., Syaputra, Y. D., Ummah, Y. K., and Muhamad, A. F. (2023). Religious moderation scale (rm-s): validitas, reliabilitas dan implikasinya dalam bimbingan dan konseling. *Edu Consilium : Jurnal Bimbingan dan Konseling Pendidikan Islam*, *4*(1), Article 1.

https://doi.org/10.19105/ec.v4i1.7939

Mudanta, K. A., Astawan, I. G., and Jayanta, I. N. L. (2020). Instrumen penilaian motivasi belajar dan hasil belajar IPA siswa kelas V Sekolah Dasar. *Mimbar Ilmu*, *25*(2), 101. https://doi.org/10.23887/mi.v25i2.26611

Novitasari, A., Ridlo, S., and Kristina, T. N. (2017). Instrumen Penilaian diri kompetensi klinis mahasiswa kedokteran. *Journal of Educational Research and Evaluation*, *6*(1), 81–89. https://doi.org/10.15294/jrer.v6i1.16212

Nurislami, I., Suarni, N. K., and Suranata, K. (2021). Pengembangan instrumen pengukuran karakter self awareness pada masa pra remaja. *Jurnal Bimbingan Dan Konseling Indonesia*, *6*(1), Article 1. https://doi.org/10.23887/jurnal_bk.v6i1.597

Nuryana, M. B., Fatimah, S., and Yuliani, W. (2022). Validitas dan reliabiltas angket prokrastinasi akademik. *FOKUS : Kajian Bimbingan dan Konseling Dalam Pendidikan*, *5*(5), 352–357. https://doi.org/10.22460/fokus.v5i5.7409

Ramdani, D., Supriatna, E., and Yuliani, W. (2023). Validitas dan reliabilitas angket kematangan emosi. *FOKUS : Kajian Bimbingan dan Konseling Dalam Pendidikan*, *6*(3), 232–238. https://doi.org/10.22460/fokus.v6i3.10869

Rozzaqyah, F., AR, S., and Wisma, N. (2021). Pengembangan inventori kecenderungan perilaku agresif siswa SMP. *Jurnal Bimbingan dan Konseling Ar-Rahman*, *7*(1). https://ojs.uniska-bjm.ac.id/index.php/BKA/article/view/4463/3054

Suradipa, I. (2022). Pengembangan Instrumen literacy humanistic dan kemampuan berpikir kritis pada pembelajaran IPA kelas V SD. *Pondasi Jurnal Pendidikan Dasar Indonesia*, *6*(1).

Sürücü, L., and Maslakçi, A. (2020). Validity and reliability in quantitative research. *Business and Management Studies: An International Journal*, *8*(3), 2694–2726. https://doi.org/10.15295/bmij.v8i3.1540

Suryani, N. (2021). Validitas dan reliabilitas instrumen korban bullying. *Jurnal Hisbah*, *18*(2), 192-204.

Taopik, I., Supriatna, E., and Yuliani, W. (2023). Uji validitas dan reliabilitas angket interaksi sosial. *FOKUS : Kajian Bimbingan dan Konseling dalam Pendidikan*, *6*(4), 278–284. https://doi.org/10.22460/fokus.v6i4.11060

Triandini, E., Jayanatha, S., Indrawan, A., Putra, G. W., and Iswara, B. (2019). Metode systematic literature review untuk identifikasi platform dan metode pengembangan sistem informasi di Indonesia. *Indonesian Journal of Information Systems (IJIS)*, *1*(2).

Tussaadah, R. Z., Hendriana, H., and Yuliani, W. (2021). Validitas dan reliabilitas angket penyesuaian diri peserta didik SMA. *FOKUS : Kajian Bimbingan Dan Konseling Dalam Pendidikan*, *4*(3), 213–218. https://doi.org/10.22460/fokus.v4i3.7250

Wulandari, A., and Radia, E. (2021). Pengembangan instrumen penilaian sikap tanggung jawab pembelajaran tematik terpadu kelas V SD. *MIMBAR PGSD Undiksha*, *9*(1), 10. https://doi.org/10.23887/jjpgsd.v9i1.32979

Zhang, Y., Huang, Y., Qiu, Y., Hu, Y., Tao, J., Xu, Q., Zhang, L., and Dong, C. (2024). Trajectories and influencing factors of psychological resilience among Chinese patients with

maintenance hemodialysis. *Heliyon*, *10*(17), e37200. https://doi.org/10.1016/j.heliyon.2024.e37200

Zikra, Z., and Laras, N. P. (2021). Analisis soal ulangan materi pemanasan global untuk SMP. *Jurnal Bioconcetta*, *6*(2), 49–56. https://doi.org/10.22202/bc.2020.v6i2.4175