

A Revision Process That Bridges Qualitative and Quantitative Assessment

Louise Stroud^{1,2} , Elizabeth Green^{1,2} , Johan Cronje¹ 

¹Nelson Mandela University, Port Elizabeth, South Africa

²Association for Research in Infant and Child Development, Manchester, UK

Email: louise.stroud@mandela.ac.za

How to cite this paper: Stroud, L., Green, E., & Cronje, J. (2020). A Revision Process That Bridges Qualitative and Quantitative Assessment. *Psychology, 11*, 436-444.
<https://doi.org/10.4236/psych.2020.113029>

Received: February 12, 2020

Accepted: March 14, 2020

Published: March 17, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

There is an increasing trend to consider complementary ways of combining quantitative and qualitative assessment. Furthermore, there is a growing awareness of the need to revise psychological measures on a regular basis. This article attempts to address both these trends and uses a child development measure, the Griffiths III, as an illustration of how this can be accomplished. The outcome proposes a potential six step revision process.

Keywords

Child Development, Test Revision, Quantitative Measurement, Qualitative Measurement, Griffiths III

1. Introduction

Not everything in life becomes functionally ineffective at the same rate; this applies to psychological measures some of which become dated and require more frequent or more extensive revision over time than others. Revisions of psychological measures are a historical fact as it is well known that they are susceptible to the effects of time (Aylward & Aylward, 2011; Butcher, 2000). In particular, it is important for measures of child development to be updated regularly as they need to accurately reflect stages of development in the present times and to maintain links with a changing broader educational context.

When is a psychological measure in need of revision? Changing from an established and relied-on measure to an improved version is often difficult for clinicians and few adapt to a revision without a struggle. There is a considerable amount of literature that concerns procedures for the development and revision of quantitative instruments (e.g., Coaley, 2010; Hogan, 2013; Luyt, 2012; Rust &

Golombok, 2008). Yet an integrated framework for the revision of psychological measures, in which both qualitative and quantitative methods of clinical interpretation and application of the data obtained are incorporated and provided for, appears lacking.

Both qualitative and quantitative data can be regarded as “thin descriptions” and both need an interpretative act to become “thick descriptions” (Love, 2013). Hence, the impetus for this conceptual piece of work arose from an identified need informed to a large extent by practice. The Griffiths Scales of Child Development (Griffiths III; Stroud et al., 2016), which was originally developed in 1954, is a child development assessment measure that has already undergone a number of revision iterations since its initial development. Used as an illustrative case study in this article, the revision process of the Griffiths III demonstrates how qualitative and quantitative data may be mixed so that “thin descriptions” of child development are turned into “thick descriptions”.

2. Context

There has been widespread debate in recent years regarding the relative merits of implementing quantitative and qualitative strategies for the interpretation and application of data or information obtained from the administration of psychological measures. The position taken by clinicians varies considerably, from those who see the two strategies as entirely separate and based on alternative views to those who are happy to integrate them. Regardless of strategy, the appropriate interpretation of the data obtained in the assessment of children is vital, especially as child development is known to be a dynamic, transactional process. As such it should be possible for the clinician to combine quantitative and qualitative assessment data with confidence.

A mixed methods approach of this kind offers several advantages. This includes the acknowledged observation that the combined use of quantitative and qualitative approaches provides a better understanding of the problems than either approach alone (Creswell & Plano Clark, 2018). There is a growing belief among child development specialists that such combined data obtained in an assessment of a child can offer a firmer foundation for the decisions needing to be made regarding their development. To this end, the current article describes the manner in which the revision of a child development measure such as the Griffiths III (Stroud, et al., 2016) envisages the complementary incorporation of qualitative and quantitative data that could be obtained in the assessment of a child.

3. Revising a Child Development Assessment Measure

Psychological measures are revised for a number of reasons. Among these are to: update test norms and new psychometric information (e.g., reliability coefficients for using a test in a clinical or cross-cultural settings); revise the mode of administration or response required in the light of advances in test development

and adaptation; evaluate the relevance of the construct(s) tapped and the construct coverage in relation to ongoing research of the construct; and refresh the test items (e.g., Bush, 2010; Strauss, Spreen, & Hunter, 2000). Interestingly, the age of the measure, in isolation, is not a significant determinant in making the judgment as to whether to revise.

As psychologists and clinicians need to select measures from among several available options, maintaining test user confidence is critical to the success of the revision of a measure. The revision process of measures takes various forms. While opinions differ about when a measure should be revised, Adams (2000) observes that new versions of clinical measures now appear at approximately 10-year intervals, a shorter time-span than the “one generation” postulated by Silverstein and Nelson (2000). This short life span of a measure and the period required to prepare the groundwork for a revision means that a measure should generally be treated as a work in progress, with research and development featuring as a standing item on the measure’s agenda.

One stimulus that is used to motivate for a revision is the Flynn effect, named after Professor Jim Flynn (1984, 1987) who highlighted the concept in the 1980s. The Flynn effect refers to the continuous rise in observed average scores, especially intelligence tests from the 1930s onwards. The effect of this gradual rise in scores is that test norms lose validity over time and become significantly inaccurate after 10 to 15 years (Aylward & Aylward, 2011; Silverstein & Nelson, 2000). A meta-analysis conducted on 285 studies confirmed previous estimates of the Flynn effect across various ages (Trahan et al., 2014). Measures are also revised if the language used in tasks or instructions becomes outdated or confusing for modern test-takers. Task performance characteristics, such as their difficulty or discrimination level, may also change over time, thereby affecting test performance.

A further aspect that necessitates a revision is change in the constructs that underpin a measure. Any measure is grounded in theory or an understanding of the construct that underpins it. Over time, theories develop and research expands the understanding of a theoretical construct. A revision provides an ideal opportunity to revisit the foundations of a measure, to ensure its academic credibility (Silverstein & Nelson, 2000). Psychometrics is a developing field, and through revision a measure can be updated in line with recent measurement theory, such as moving from classical test theory to item response theory, or from norm-referenced testing to criterion-referenced testing.

This article highlights key aspects of the revision process illustrated by the Griffiths III. This specific revision process could not fit neatly into extant test revision descriptions and processes as the newly revised Griffiths III needed to bridge the gap between the quantitative and qualitative data obtained and, in so doing, turn “thin descriptions” of the child being assessed into “thick descriptions”. It also needed to provide an innovative approach to child developmental assessment in which the more traditional dichotomy created by quantitative and qualitative data obtained was challenged.

4. The Griffiths Scales

Ruth Griffiths designed and developed the Griffiths Mental Development Scales (Griffiths, 1954, 1970). Her philosophy of child development was based on the observation of children, the basic avenues of learning, the concept of play, and the need to assess with greater confidence the development of young children. At the time the Griffiths Mental Development Scales were first introduced, psychometric conceptions of intelligence were emerging and were to influence psychometric measurement for the next three generations. These narrow conceptions included verbal, visual-spatial and mathematical abilities. The Griffiths Mental Development Scales introduced an innovative system for developmental assessment, as Griffiths was acutely aware of the importance of interaction between the various avenues of learning.

Griffiths advocated a broad-based approach to understanding mental development (i.e., the processes and rates at which growth and maturation of a child's attributes and abilities take place). She was aware of the importance of social and emotional developmental factors and the interplay between these and mental development. Importantly, in the field of childhood development today, the areas underpinning the Griffiths Mental Development Scales remain critical aspects of child development, the combination of which provides a holistic blueprint of mental development in infancy and early childhood.

Child development is a dynamic, moving process, and questions constantly need to be asked. Therefore, it is important from time to time to take a fresh look at the underpinning theories, philosophies and principles of child development to ensure that these remain relevant. If they do not, a revision of the measure must be considered. One challenge when revising or updating a test is the thin line between modernization and the retention of the original spirit of the measure. The periodic review of assessment measures such as the Griffiths Scales is thus considered good practice to ensure the continued ethical soundness of the scales as psychometric measures.

5. Predecessors of the Griffiths III

Published in 1954, the Griffiths Scales of Mental Development covered the period from birth to two years (later known as the Baby Scales) and was standardised on a sample of 571 babies in Britain (Stewart, 2005). It assessed five avenues of learning, namely Locomotor, Personal-Social, Hearing and Speech, Eye and Hand Co-ordination, and Performance. Ruth Griffiths received requests from professionals for the extension of the infant scales for use in clinical practice with older children. Therefore, the Extended Griffiths Scales were developed in the 1960s and described in her book *The Abilities of Young Children* (Griffiths, 1970). This expanded Griffiths Scales covered from birth to eight years and a sixth subscale, Practical Reasoning, was added to the Extended Griffiths Scales for children aged two years and older.

A need for a further revision of the Griffiths Scales was suggested by various studies in the 1980s and early 1990s. In 1994 the publication of a draft version of the Revised Baby Scales from Birth to Two Years was realised. The aim was to preserve the original, whilst only making those changes necessary to update the measure and re-norm it for the present times (Huntley, 1996). This update was more consistent with a “medium” revision, i.e., it was more intensive and included changes to or replacement of non-performing tasks, and the updating of the norms of the measure (Butcher, 2000).

A decision to revise the Extended Griffiths Scales was made concurrently. This decision culminated in the publication of the Griffiths Mental Development Scales—Extended Revised (GMDS-ER) (Luiz et al., 2006). An “extensive” revision involves a complete reanalysis and reconstruction of the test. This could involve relooking at the theoretical foundation of the test, major changes to tasks or subscales, as well as a new set of test instructions. An extensive revision would also include new standardisation data, as well as validity and reliability studies. The Extended Revised Scales represents a medium revision as it only contained some elements of an extensive revision. The revision was aimed at replacing non-performing tasks, broadening the areas of development covered by new tasks, improving stimuli and instructions, and updating the norms and scoring (Luiz et al., 2006). The test developers also made use of technology by introducing instructional DVDs and computerised scoring.

The revised versions of both the Baby Scales and the GMDS-ER kept their original format and stayed true to Griffiths’ vision of creating a sound and accurate assessment measure that remains natural and non-threatening to children, whilst keeping abreast of current knowledge of the behaviour and development of young children and infants (Luiz et al., 2006). The Griffiths III, however, is an extensive revision of the Griffiths Scales. In the process of revising the Griffiths Scales, a complete re-analysis of the underpinning foundations was undertaken with the scales being developed from the construct level upwards while still preserving a quintessential “Griffiths feel”. Restandardisation was also undertaken, thus providing a statistically sound, valid and reliable measure of a child’s development with a new set of norms.

With the above considerations in mind an extensive revision of the Scales was warranted. This revision process aimed to ensure that the Griffiths Scales tapped into unique child development areas and, in so doing, be set apart from other child development measures. The decision also had to be made as to whether the revision process should focus on the European context only, or if it should be extended internationally, thus affording the Griffiths Scales greater flexibility in terms of cross-cultural use and wider relevance. This decision necessitated the exploration of possible alternative methodologies of assessing child development (with regards to globalisation, testing and standardisation), given that other major international measures had adopted a more global route in recent test development and norming processes. The revision also considered the mode of test-

ing and the context of application in the revision process. This meant keeping in mind that the Griffiths Scales is in essence a “child friendly” developmental measure, based on the skill and value of observing children, and that it is playful in nature. It is these attributes which have rendered the Scales successful in the assessment of children, especially those with clinical diagnoses who may be threatened by a formal, rigid testing situation.

At the outset of the revision, appropriate test specifications for the Griffiths III had to be established and stipulated, with the initial question needing to be asked being: “what is normal development?” As the Griffiths is based on chronological age, defining normal child development specifically in terms of developmental tasks was necessary. In other words, the concept of chronological age needed to be unpacked into different variables based on current trends in child development research and contemporary child development theory. The concept also needed to reflect sensitively where a child’s development deviates from the norm (i.e., once it has been established what the norm is). It is important to recognise that once developmental tasks have been identified and established and sensitive specificity built into the test, a balance between the two variables of typical and atypical must be achieved. This ensures that the developmental nature of the Griffiths Scales would be retained.

As the significance of understanding the process of early childhood development has been increasingly acknowledged, so too has the need to confirm the value and role of developmental assessment and the “type” of data being obtained. Hence, the importance of the decision taken to revise the Griffiths with a six phase approach or structure guiding the process. This revision process is illustrative of how a test revision can occur.

6. The Griffiths III Six Phase Revision Process

The revision of the Griffiths III was divided into six phases. At the outset a scoping of relevant and recent literature, stakeholder feedback and market research was undertaken. Thereafter constructs were developed and potential tasks reviewed and selected by a team of Griffiths expert psychologists and paediatricians. This was followed by experimental and pilot testing, standardisation and norm development and, finally, training that is ongoing internationally. These six phases are descriptively elaborated on in **Figure 1**.

As can be seen in **Figure 1**, a unique method or procedure was followed in the revision of the Griffiths III. Both qualitative and quantitative data were collected, analysed, and mixed at various points during the six phase revision process. The dataset which was obtained guided the revision process. The data were merged or converged so that they were able to build on each other. It was also possible to embed one dataset within the other so that one type of data provided a supportive role for the other dataset (Creswell & Plano Clark, 2018). This approach allowed for the validation of the quantitative statistical findings with the qualitative data results.

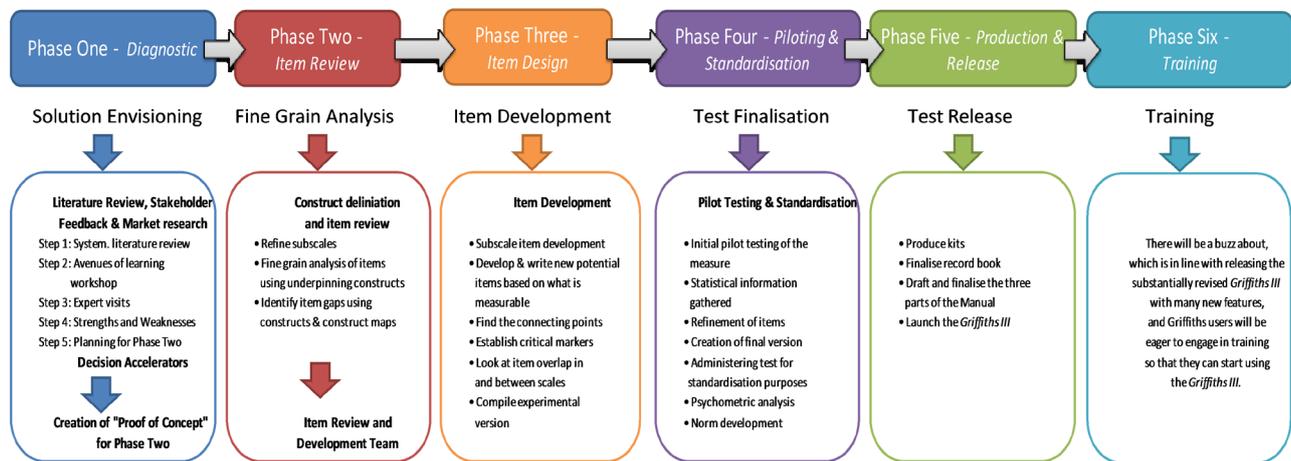


Figure 1. The Griffiths III six phase revision process (Stroud et al., 2016: p. 7).

This complementarity of the qualitative and quantitative dataset provided a rich elaborated understanding and placed equal value on both consistent and inconsistent findings obtained. Mixing qualitative and quantitative data is considered beneficial because the two forms of data are able to capture multiple realities and new explanations and decisions are able to emerge as a result (Luyt, 2012; Ponterotto & Grieger, 1999). In this sense, neither dataset is superior to the other in describing a single aspect of child development. Rather, both contribute toward the understanding, description, and exploration of multiple aspects of children’s development.

The clinician who can “wear two hats” so to speak, shifting in a sequenced and integrative fashion between descriptive and normative approaches, will be more effective and better able to capture the true complexity of a child’s development. Psychological assessment engages the same duality, the distinction between the latent psychological characteristic and particular empirical phenomena used to gauge that characteristic (Knowles & Condon, 2000). The revised Griffiths III provides the clinician with a child assessment measure that, in its application, is able to yield information far beyond the assumptions made using psychometric or qualitative data on their own. In other words, the divide between these two distinct datasets becomes integrated, and the “thin” descriptions of the child’s development become usefully “thick” descriptions.

7. Conclusion

This article suggests an effective process for merging qualitative and quantitative findings in measurement revision through a six phased procedure or approach. Neither quantitative nor qualitative data were considered merely supplementary as both provided unique as well as supplementary information in the revision process. As such equal status should be granted in terms of the sum of their individual contributions.

The Griffiths III, presented as an illustrative case study in this article, offers an example of how qualitative and quantitative data can be used in a revision process.

This in turn provides a framework for the interpretation and application of the data obtained when a measure such as the Griffiths III is used to assess a child's development. Ultimately, "thin descriptions" of child development can be turned into "thick descriptions" in the revision process suggested. It is hoped that this revision process remains sufficiently flexible so that it can be applied to a range of different cases.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Adams, K. M. (2000). Practical and Ethical Issues Pertaining to Test Revisions. *Psychological Assessment, 12*, 281-286. <https://doi.org/10.1037/1040-3590.12.3.281>
- Aylward, G. P., & Aylward, B. (2011). The Changing Yardstick in Measurement of Cognitive Abilities in Infancy. *Journal of Developmental and Behavioural Pediatrics, 32*, 465-468. <https://doi.org/10.1097/DBP.0b013e3182202eb3>
- Bush, S. S. (2010). Determining Whether or When to Adopt New Versions of Psychological and Neuropsychological Tests: Ethical and Professional Considerations. *The Clinical Neuropsychologist, 24*, 7-16. <https://doi.org/10.1080/13854040903313589>
- Butcher, J. N. (2000). Revising Psychological Tests: Lessons Learned from the Revision of the MMPI. *Psychological Assessment, 12*, 263-271. <https://doi.org/10.1037/1040-3590.12.3.263>
- Coaley, K. (2010). *An Introduction to Psychological Assessment and Psychometrics*. London: SAGE.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). Thousand Oaks, CA: SAGE.
- Flynn, J. R. (1984). The Mean IQ of America: Massive Gains 1932 to 1978. *Psychological Bulletin, 95*, 29-51. <https://doi.org/10.1037/0033-2909.95.1.29>
- Flynn, J. R. (1987). Massive IQ Gains in 14 Nations: What IQ Tests Really Measure. *Psychological Bulletin, 101*, 101-191. <https://doi.org/10.1037/0033-2909.101.2.171>
- Griffiths, R. (1954). *The Abilities of Babies*. London: University of London Press.
- Griffiths, R. (1970). *The Abilities of Young Children, 0-8 Years*. London: Association for Research in Infant and Child Development (ARICD).
- Hogan, T. P. (2013). *Psychological Testing: A Practical Introduction* (3rd ed.). New York: John Wiley.
- Huntley, M. (1996). *The Griffiths Mental Development Scales. From Birth to 2 Years. Manual* (Revision). London: ARICD.
- Knowles, E. S., & Condon, C. A. (2000). Does the Rose Still Smell as Sweet? Item Variability across Test Forms and Revisions. *Psychological Assessment, 12*, 245-252. <https://doi.org/10.1037/1040-3590.12.3.245>
- Love, H. (2013). Close Reading and Thin Description. *Public Culture, 25*, 401-434. <https://doi.org/10.1215/08992363-2144688>
- Luiz, D. M., Faragher, B., Barnard, A., Knoesen, N., Kotras, N., Burns, L. E., & Challis, D. (2006). *GMDS-ER: Griffiths Mental Development Scales—Extended Revised Analysis Manual*. Oxford: Hogrefe—The Test Agency Ltd.

- Luyt, R. (2012). A Framework for Mixing Methods in Quantitative Measurement Development, Validation, and Revision: A Case Study. *Journal of Mixed Methods Research*, 6, 294-216. <https://doi.org/10.1177/1558689811427912>
- Ponterotto, J. G., & Grieger, I. (1999). Merging Qualitative and Quantitative Perspectives in a Research Identity. In M. Kopala, & L. A. Suzuki (Eds.), *Using Qualitative Methods in Psychology* (pp. 49-62). Thousand Oaks, CA: Sage Publications, Inc. <https://doi.org/10.4135/9781452225487.n5>
- Rust, J., & Golombok, S. (2008). *Modern Psychometrics: The Science of Psychological Assessment* (3rd ed.). New York: Routledge/Taylor & Francis Group.
- Silverstein, M. L., & Nelson, L. D. (2000). Clinical and Research Implications of Revising Psychological Tests. *Psychological Assessment*, 12, 298-303. <https://doi.org/10.1037/1040-3590.12.3.298>
- Stewart, R. (2005). *The Griffiths Scales of Mental Development: Proposed New Items for the Revision of Subscales A, B, C, D and E*. Unpublished Doctoral Dissertation, Port Elizabeth: Nelson Mandela Metropolitan University.
- Strauss, E., Spreen, & Hunter, M. (2000). Implications of Test Revisions for Research. *Psychological Assessment*, 12, 237-244. <https://doi.org/10.1037/1040-3590.12.3.237>
- Stroud, L., Foxcroft, C., Green, E., Bloomfield, S., Cronje, J., Hurter, K., Venter, D. et al. (2016). *Griffiths Scales of Child Development 3rd Edition; Part I: Overview, Development and Psychometric Properties*. Oxford: Hogrefe.
- Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The Flynn Effect: A Meta-Analysis. *Psychological Bulletin*, 140, 1332-1360. <https://doi.org/10.1037/a0037173>