Visual scaling to Improve Organizational Decisions in Educator Preparation: Effectively Comparing Apples and Oranges

Raymond W Francis

Central Michigan University

Mark E Deschaine

Central Michigan University

Author Contact

Dr. Raymond W Francis

447 EHS Building, Central Michigan University, Mt Pleasant MI 48859

franc1rw@cmich.edu

February 19, 2018

A Paper Presented at the 2018 Association of Teacher Educators Annual Meeting in Las Vegas, Nevada. February 17 – 20, 2018

Keywords: Visual Scaling, data-based decision making, program evaluation, assessment

Table of Contents

Introduction: The need for Comparisons Between and Among Data

One of the largest challenges in educator preparation in higher education is the effective use of data for actionable and meaningful decision making about programs for current and future educators, administrators, and educational leaders. Within programs there are many types of students. These may include elementary, secondary, special education, foreign language, social studies, science, literacy, health, school level administrators, district administrators, curriculum leaders, and many others based on school and program. However, in all instances the program (i.e. college, university, school, etc.) is charged with using data in a meaningful manner to guide program decisions, make revisions to curriculum and clinical experiences, and determine the overall quality of the program. Visual Scaling provides an effective solution, and is a process whereby goals and expectations are clearly articulated, and baseline and performance assessments are compared against the expected level of performance.

This is where a very simple, yet extraordinarily effective process comes into play. That process is call Visual Scaling (VS), and the purpose of Visual Scaling is to allow for the comparison of collected data from which decisions can effectively be made. In Visual Scaling identified goals and expectations are clearly articulated, and baseline and performance assessments are compared against the expected level of performance. This process then allows data to be compared within and between representative groups, samples, and populations. As a result, effective and meaningful data-based decisions can be made and educator preparation programs can enhance their assessment/evaluation and accreditation processes to best serve their candidates at all levels.

Applied and Theoretical Perspective

History in Clinical Applications

The foundation of scaling, in general, emerges from the medical field where practitioners sought to establish comparative scales of client/patient performance to determine progress in rehabilitation and recovery programs (Bouwens, van Heugten, & Verhey, 2009). Practitioners were able to establish overall health and a wellness scales to use in exploring patient rehabilitation and growth (Rockwood, Graham, & Fay, 2010). The ideas of effective and comparative measurement were attractive to the health care community (Terwee, et al., 2010), and the transferability of the concept of scaling became accepted in clinical settings, as well as settings outside of the health care field. As clinical successes led to the exploration of the use of scaling in areas outside of health care, the need for a practical application of data from evaluation processes in educator preparation became an imperative.

Over time practitioners have honed and refined their skills and instruments (Mokkink, et al., 2010) to enable more consistent and effective communication of knowledge in patient care (Moher, et al., 2010). The ongoing refinement of the use of scales continues to improve the quality and effectiveness of patient care (Terwee, 2007).

Overall, the importance of the use of the scaling process became the consistently resetting of the base-zero, foundation data. As patents' conditions improved, or became worse, the base-zero of the individual patient could effectively be referenced and identified to be within the expected, or non-expected, zones to allow for the revision of patient treatment plans.

Scaling in Traditional Perspectives

Standard perceptions of comparative data provided the opportunity for data comparison across program types, levels, and certifications (Fitzpatrick, Sanders, & Worthen, 2004). This was provided as a result of the general acceptance of standards-based evaluation, criteria- based assessment, and the need for larger sample size in programs with small numbers of future

educators (Phillips & Phillips, 2016). Scaling, as a program evaluation, accreditation, and assessment process allows for the comparison of unique data (Coombs, 1964), allows for the amplification of n values in computation, provides for the comparison of data (Carlier, Lamberts, & Gersons, 2000), and provides for effective data in the assessment and accreditation process (Gardner & Gardner, 2012). In addition, issues related to internal and external validity, reliability, trust, comparison, and interpretation are all effectively addressed within and between comparison groups. However, in general, these traditional perspectives of the use of the concept of scaling were limited and provided for only a small entre into the idea of comparative data in evaluation.

Need for Visual Scaling in Higher Education

Scaling, as a program evaluation, accreditation, and assessment process allows for the comparison of unique data (Coombs, 1964), allows for the amplification of n values in computation, provides for the comparison of data (Carlier, Lamberts, & Gersons, 2000), and provides for effective data in the assessment and accreditation process (Gardner & Gardner, 2012). In addition, issues related to internal and external validity, reliability, trust, comparison, and interpretation are all effectively addressed within and between comparison groups.

Implications for Policy

Visual Scaling is a practical tool for data-based processes and evaluation/assessment within and between programs. Scaling provides the opportunity for institutions to explore their own policy and practices to determine effective practices related to evaluation/assessment within programs, and about candidates. Visual Scaling, as presented, also provides the opportunity for institutions to establish policy and practice and define their own expectations as a part of evaluation/assessment and accreditation processes

The Visual Scaling (VS) Process

A key aspect of the VS process is the idea of comparison of data to an identified or expected result, or an identified or expected standard. The data is being collected from a variety of observational, survey, performance, rating, or other instruments. However, each set of collected data is compared to an expected result or expected standard. That expected result or expected standard represents the zero-base level of the process from which all other data, results, or performance results is compared. From the zero-base level, other levels, with identified criteria and comparative

Identify Expected Zero-Based Level of Performance

The foundation of the VS process is the identification of an expected level of performance, or expected zero-base level. This expected zero-base level of performance is "the initial expected level of participant performance for an identified event" (CITATION, DATE). This identified level of expected performance may be differentiated for segments of a population and for tasks. In some cases, without prior knowledge or experiences, the initial expected zero-base level of performance is developed as a result of the professional knowledge base, prior experiences, or ongoing face validity established by the program. In addition to establishing the expected zero-base level, the program must also establish the acceptable range for the zero-base range, as well as the ranges for the +1, -1, +2, -2, +3, and -3 levels.

These identified ranges and levels are not intended to be statistically based, or based on ideas such as Gaussian distributions, standard deviations, or similar concepts. The ranges and levels are intended to represent realistic and meaningful ranges and levels for the program that represent participant performance, and are based on professional expectations of performance.

One possibility in establishing the expected zero-base level of performance would certainly be to make use of existing descriptive and inferential calculations for items such as

mean, mean, mode, standard deviation, correlation coefficient, and other features. The result of this established process would be to inform the program about the participant performance on a scale and range that, although statistically appropriate, these results may be meaningless to the program.

Consider the option of using a scale from -3 to 0 to -3. For example, in a situation where the expected zero-base performance level is established by the program to be 64 to 86 (arbitrary for purposes of the example), and the other ranges are determined in a similar manner. This would allow for any performance score in this range to be converted to the Visual Scaling range.

+3	=	98 – 100
+2	=	95 – 97
+1	=	87 – 94
0 Zero-base	=	64 – 86
-1	=	48 – 63
-2	=	18 – 47
-3	=	0 – 17

For our purposes in this work, the scale is an irregular scale with irregular intervals. However, it represents an actual scale that might be developed by a program that has considered participant's past performance, difficulty of the performance event, professional expectations within the program, and known levels of expected participant experience. This type of scale provides for the program to be specific about expected levels of performance in the initial stage of evaluation, and provides for participants to be informed of the expected levels of performance. The identified levels can be differentiated based on tasks, student level in the program, etc. to enable to effective collection of data.

Identify Change in Level of Performance over Time

The next step in the Visual Scaling process is to collect data from the participants and explore the relationship of the participant data to the expected zero-base level of performance.

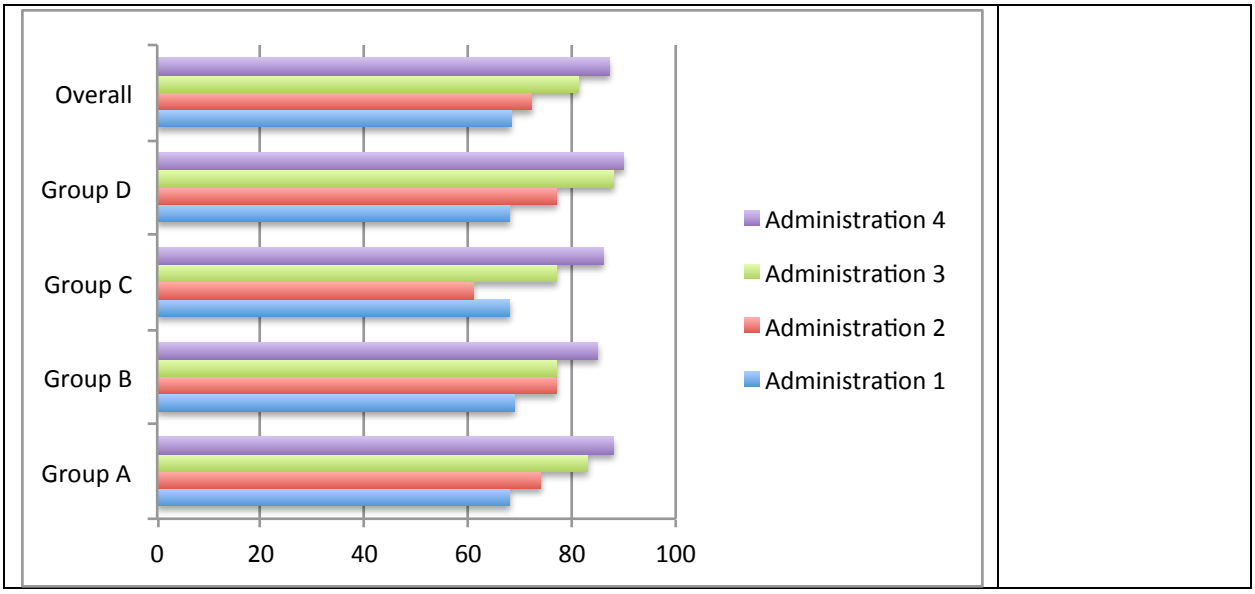
+3	=	98 – 100
+2	=	95 – 97
+1	=	87 – 94
0 Zero base	=	64 – 86
-1	=	48 – 63
-2	=	18 – 47
-3	=	0 – 17

Table 1: Sample of 4 Administration Data for Heterogeneous Group over Time

	Administration 1	Administration 2	Administration 3	Administration 4
Group A	68	74	83	88
Group B	69	77	77	85
Group C	68	61	77	86
Group D	68	77	88	90
Overall	68.25	72.25	81.25	87.25

From a traditional perspective, we can see that the overall scores for the administration of an intervention over time produce slight increases as an overall perspective. In short, each administration of the assessment demonstrates an overall increase when compared to the previous data collected.

Chart 1: Sample of 4 Administration Data for Heterogeneous Group over Time



In short, our interpretation of Table 1 is supported through our review of Chart 1 where we see the overall performance of each group through the 4 administrations of an assessment is a relatively positive graph with a similar format and pattern.

However, when we employ a Visual Scaling process, we see a very different process. For example, overall in Table 2, we see that most of the groups and administrations of the assessment are demonstrating a zero-base performance. In other words, almost all of the administrations of the assessment demonstrate that participants are performing within the expected levels. Yet, some of the groups (Group C - #2, Group D - #3, Group D - #4, and Overall - #4) are at a performance level outside of the expected performance levels for the participants. In the case of Group C -#2, the data indicates the group is performing below the expected level during Administration #2. However, Group C recovers during Administration #3 and Administration #4. In addition, in the Overall (total) group the performance level is above expectations.

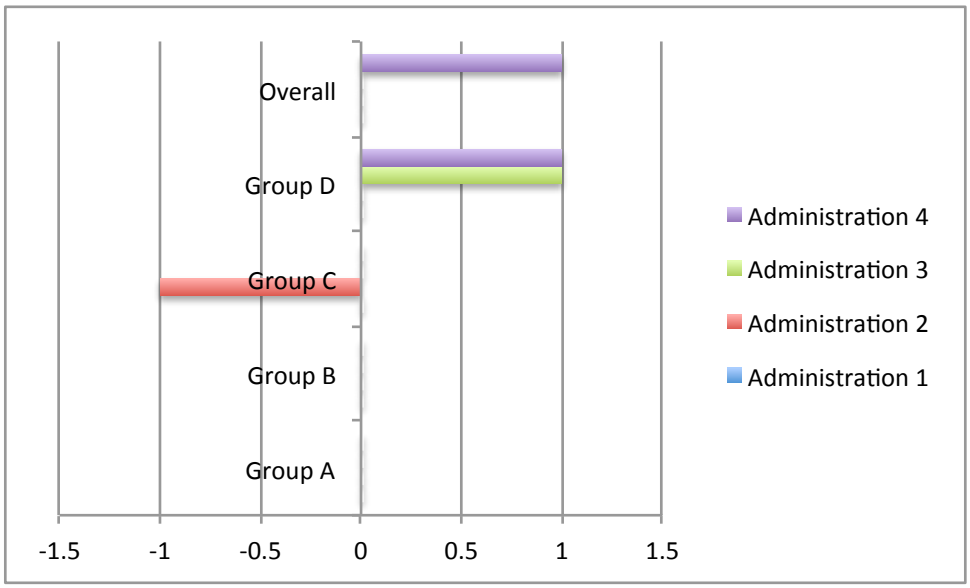
Table 2: VS Results for a Sample of 4 Administration Data for Heterogeneous Group over Time

	Administration 1	Administration 2	Administration 3	Administration 4
Group A	0	0	0	0
Group B	0	0	0	0
Group C	0	-1	0	0
Group D	0	0	+1	+1
Overall	0	0	0	+1

The results from Table 2 are easily verified through examination of the results presented in Chart 2. As presented in Table 2, Groups A and B performed at the expected level throughout the process. While group C performed below expectation during only Administration 2. Group D demonstrated a level of performance above the expected level during Administration 4, and as an

overall performance level. Lastly, the Overall result demonstrated a performance level above expectation for the entire process.

Chart 2: VS Results for a Sample of 4 administration data for heterogeneous group over time



Implications for Practice

Future Actions

Overall, the process of Visual Scaling has great potential in educator preparation at all levels. The process can be used as a specific inter-program or intra-program process for making effective data-based decisions related to program review and revisions. In addition, the Visual Scaling process can be used to review and evaluate individual and group progress over time as a result of experience, interventions, course work, or other individual learning factors.

The future actions connected to the Visual Scaling process include an expansion of the exploration of the types and forms of data that can be compared and reviewed within the Visual Scaling process. These data include school district, grade level assessments, high stakes state

assessments, and many other school-based data. The Visual Scaling process also has implications across the realm of higher education through providing an avenue to compare and contrast programs on a level playing field with meaningful and actionable date.

More work is needed, and the results will be shared.

Bibliography

Bouwens, S. F. M. , van Heughten, C. M., & Verhey F. R. J. (2009). The practical use of goal attainment scaling for people with acquired brain injury who receive cognitive rehabilitation. *Clinical Rehabilitation*, *23*, 310-320.

Coombs, C. H. (1964). *A theory of data*. CITY, ST: PUBLISHER.

Carlier, I. V., Lamberts, R. D., & Gersons, B. P. (2000). The dimensionality of trauma: A multidimensional scaling comparison of police officers with and without posttraumatic stress disorder. *Psychiatry Research*, *97*(1), 29-39.

Gardner, J. N., & Gardner, J. (Eds.). (2012). *Assessment and learning*. CITY, ST: Sage.

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2004). Program evaluation: Alternative approaches and practical guidelines. CITY, ST: PUBLISHER.

Kiresuk, T. J., Smith A., & Cardillo, J. E. (2014). *Goal attainment scaling: Applications, theory, and measurement*. London, England: Psychology Press.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group P. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, *8*(5), 336–341.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., et al. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Resources*, *19*(4):539–549.

Phillips, J. J., & Phillips, P. P. (2016). *Handbook of training evaluation and measurement methods*. CITY, ST: Routledge.

Rockwood, K, Graham, J. E., Fay, S. (2002). Investigators: A goal setting and attainment in Alzheimer's disease patients treated with donepezil. *Journal of Neurological Neurosurgical Psychiatry*, *73*(5), 500–507.

Terwee, C., Mokkink, B., Poppel, L., Chinapaw, M., Mechelen, N., & Vet, M. (2010). Qualitative attributes and measurement properties of physical activity questionnaires. *Sports Medicine, 40*(7), 525-537.

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker J, et al. (2007). Quality criteria were .proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34–42.