



## The Journal of Sociology & Social Welfare

Volume 24  
Issue 4 *December*


Article 2

December 1997

# The Biography of a Scale: Contextual Factors That Influence the Measurement of Family Functioning

Ludwig Geismar  
*Rutgers University*

Follow this and additional works at: <https://scholarworks.wmich.edu/jssw>

 Part of the [Marriage and Family Therapy and Counseling Commons](#), and the [Social Work Commons](#)

### Recommended Citation

Geismar, Ludwig (1997) "The Biography of a Scale: Contextual Factors That Influence the Measurement of Family Functioning," *The Journal of Sociology & Social Welfare*: Vol. 24 : Iss. 4 , Article 2.  
Available at: <https://scholarworks.wmich.edu/jssw/vol24/iss4/2>

This Article is brought to you for free and open access by the Social Work at ScholarWorks at WMU. For more information, please contact [maira.bundza@wmich.edu](mailto:maira.bundza@wmich.edu).



# The Biography of a Scale: Contextual Factors That Influence the Measurement of Family Functioning

LUDWIG GEISMAR

Rutgers University  
School of Social Work

*The subject of instrument relevance is addressed by examining issues that have arisen in the use of a single scale over a forty year period. The issues revolve around the impact of varying social conditions, changing ethos, differential respondent receptiveness, and evolving research technology. Extended use of an instrument, it is argued, yields information that transcends the conventional techniques for testing instrument adequacy. The lack of opportunities in social work for accessing information on extended use of measurement tools is due, among other factors, to a preoccupation with working on subjects that are new and original and to a lack of coordination among research endeavors.*

In my work as a researcher and instructor of social work students I have often been asked whether certain instruments of social measurement should or should not be used. When coming from someone who has just completed an introductory course, the inquiry usually focuses on the reliability and validity of an instrument. That is, does the instrument have the capacity to measure relevantly and consistently the subject it set out to measure. What is often forgotten in this search for "perfect instruments" is the fact that available indices of reliability and validity are context specific. They yield evidence on instrument quality relative to given populations, social conditions, economic circumstances, and environmental factors.

The judicious advice to the prospective user of an instrument is generally to test independently for reliability and validity. This is more easily said than done, given the large investment in time

and resources in creating a useful instrument in the first place. The less desirable alternative is to justify the use of an existing tool relative to its prior performance in other, hopefully similar, contexts. Taking this route leaves unresolved the question of suitability of the instrument for the study at hand.

The present essay addresses the issue of instrument relevance from a descriptive, historical perspective by focusing on a single instrument that has been in use for nearly four decades. The goal is to illustrate the emergence of problems arising from changes in the social context and how these changes necessitate periodic revisions to assure continued instrument relevance over time and space. One of the assumptions underlying this approach is that problems exemplified by the longitudinal case approach parallel essentially the type of issues the researcher encounters in different settings at any one point in time. The single case analysis has the advantage of conveying the issues descriptively, thereby making them available to social work journal readers, perhaps a majority of them, who eschew writings employing quantitative analysis. I am assuming furthermore that the subject of instrument relevance for assessing social work intervention is significant not only to the researcher but also to every practitioner with an interest in service effectiveness.

#### The Family Functioning Scale as a Case Example

One's ability to view an instrument from a long term perspective presupposes that the instrument has been around and used for an extended period of time. Social work measurement tools, by and large, have a short tenure. The reason for this is the general absence of empirical research by the field of practice (Task Force on Social Work Research, 1992, pp.1-16). Questionnaires, scales, schedules, and related forms of data collection, more often than not, are produced in academic settings or under academic sponsorship and are used in specially financed projects (grants from federal or regional authorities and private foundations). In the absence of replication research in social work, which is not favored by sponsors, research instruments enjoy limited, spotty use in additional studies - often in pursuit of someone's advanced degree- but are not embraced by the field of practice in the quest for empirical knowledge.

The Family Functioning Scale, also known as the St. Paul Scale of Family Functioning, was developed in 1958 (Geismar and Ayres, 1959; 1960) in an effort to assess changes in the performance of psychosocial roles of dysfunctional families who received family centered treatment in St. Paul, Minnesota (Overton and Tinker and Associates, 1957). The instrument has undergone several revisions, the most recent dated 1993 (Geismar and Camasso, 1993). A review of the literature in the mid 1950s on family measurement of treatment outcome revealed few endeavors relevant to the task at hand. The single social work focused instrument with acceptable reliability and validity was the Hunt-Kogan Scale (1950) developed by researchers of the Community Service Society of New York. That scale, however, measured change in individuals rather than families, and family assessment, to the extent that it was carried out at all, required aggregating the scores of individual family members.

The theoretical approach to the family centered project, however, postulated that the family is not merely an aggregate of individuals but an entity which needs to be conceptualized as a social system (the time was the mid 1950s). Its components are social roles and social functions which tend to revolve around socially expected goals of family members. Given this framework it was possible to evaluate family roles and functions by the degree to which they make for the well-being of individual family members, the family as a group, and the surrounding community.

It required a two year effort to translate the concepts into concrete indices of behavior, a process known to researchers as operationalizing. The latter took two forms: 1. spelling out what behaviors are comprised under family roles and functions and 2. delimiting the degree to which these behaviors are in line with or contrary to family and societal well-being.

The Family Functioning Scale was a joint effort of caseworkers and researchers, based on theoretical formulations that guided the day-by-day operations of the practitioner. The Scale was therefore not a product superimposed from the outside but an instrument that utilized, with minor modifications, the thinking and language of the caseworker in the field.

### Instrument Format

Given the perspective of the Project's casework in the 1950s, information grouped in nine areas (and 27 sub-categories) of family functioning was collected by means of open-ended interviewing. The areas represented essentially three spheres of behavior: individual and role functioning, expressive functioning (focusing on interpersonal relationships) and instrumental functioning (mainly concerned with maintaining the family as a physical system by providing income, housing and health services).

The narrative of documented social functioning, the product of the interview, was rated against three basic criteria of adequacy. Was the functioning conducive to individual and family well-being, was it in line with social expectations including laws and mores of the community, and did it generate personal satisfaction and a sense of fulfillment? (for more detail see Geismar 1971).

The first version of the Family Functioning Scale was a summated rating scale, meaning a scale composed of a number of items whose values are summed to yield a single score. The open-ended interview approach was chosen to maximize respondents' motivation to talk about what for most of them represented a sensitive subject. From the very beginning the investigators took it upon themselves to address questions of reliability and validity. The first was dealt with by means of test-retest procedure, a comparison of independent ratings of interviews (Geismar and Ayres, 1959); the second was investigated with the aid of Guttman scaling (Geismar, La Sorte, and Ayres, 1962), a method that seeks to ascertain the interrelationship of scale categories by their adherence to a model of unidimensional scaling.

The attainment of satisfactory results in these tests did not make the Family Functioning Scale a universal tool. Issues arose in subsequent applications that brought into question the appropriateness of the original instrument at different times and in different settings. The issues were not necessarily related to flagging results in quantitative measures of reliability and validity. More often than not they arose in anticipation of problems tentatively identified in interviews with respondents, discussion with new research collaborators, and exposure to the new literature dealing with practice theory as well as research methodology. The present

paper presents a sampling of situations that brought to the fore the need to modify the instrument under consideration. They are discussed under the heading of varying social conditions, change in ethos, differential respondent receptiveness, and change in research technology.

### Varying Social Conditions

Social measurement is by definition contextual. It does not occur in a vacuum but assigns a value to a given phenomenon, a type of behavior, an act, a transaction, an experience, an attitude, an emotional response, etc. in the context of norms for each phenomenon to be measured.

Measurement of family functioning is normative in the sense that different types of functioning conceptualized by the scale need to be evaluated by the standards of adequacy detailed above. While scale building of this nature can not proceed without some generalizing about group norms, variations in such norms, nonetheless, occur from place to place and over time, making it necessary to carry out scale revisions.

In the case of the Family Functioning Scale we discovered, for example, that housing adequacy in the United States included among other factors ample bedroom accommodations for all family members, in spaces separated from daytime quarters. Our research in Sweden, by contrast, showed a lesser concern with separate spaces, given the acute housing shortage in urban areas at the time of the study. By contrast Sweden put a greater emphasis than did the United States on adherence to building codes. Lower adequacy ratings in Sweden would more likely indicate residence in sub-standard housing while in the American research a reduced housing score was more apt to be accounted for by crowded housing conditions.

Another contrast affecting the two countries -at least in the 1960s and before the American thrust to combat child abuse- was the use of physical force to control child behavior. Swedish legislation drew a clear line, banning physical punishment of children. The United States, by contrast, has been slow in evolving a code that will limit the physical punishment of children. Inevitably ratings of child care in the two countries had to take

into account the legislative provisions as well as the prevalent norms of behavior at specified points in time.

### Changing Ethos

In the mid-1950s when we began formulating the scale categories and scale items, out-of-wedlock births, though on the increase, carried considerable social stigma in the eyes of society as a whole as well as the agencies rendering services. In our assessment of social functioning the issue was not so much the weighing of the moral meaning of having children out-of-wedlock as assessing the social consequences of having a child viewed as illegitimate. For that reason raters were instructed to assign values "below adequate" to extra-nuptial motherhood.

In the wake of the sexual revolution of the 1960s, which had a far reaching effect on attitudes and norms of conduct, having a child out-of-wedlock was no longer seen as clearly deviant behavior, having lost the stigma attached to it in earlier decades. It seemed logical, therefore, to assess out-of-wedlock motherhood as less than adequate only in situations where it gave rise to community opposition that had a negative impact on child rearing and other aspects of family life. Again it must be noted that under the system of evaluation used, a given act or form of behavior had meaning chiefly in relation to the prevailing community norms, and this consideration shaped the ratings given to that behavior.

A similar case can be made for other non-traditional family forms such as open marriage (allowing for extra-marital liaisons) and gay marriages both of which can be evaluated only in the context of their acceptance, or rejection, by the surrounding community and the consequences attendant upon one or the other.

### Differential Respondent Receptiveness

Respondent receptiveness to study participation is a key consideration in the data collection process. Researchers who are planning to gather data are generally satisfied with a pilot study that furnishes evidence of respondents' willingness to cooperate and provide the type of data sought by the investigator. Yet, examples abound of data collection having gone awry because

changes, often minor, in type of respondents, social situation, political and economic atmosphere, interfered with obtaining the information needed for the study. The Family Functioning Scale was developed for use with clients who were given intensive family-centered services by professionally trained social workers. The semi-structured scale format, which encouraged an open-ended response, was uniquely appropriate for respondents with whom a relationship has already been established and who saw the interview as part of a helping process..

As the research expanded to include families in control and comparison groups as well as randomly selected families in a variety of social and cultural settings, it became necessary to make adjustments in the process of data gathering to achieve a measure of respondent motivation similar to that obtained from clients in service setting. In the Family Life Improvement Project (1964–1969) which encompassed 272 treatment and 283 control cases, the latter received payment for each interview. They were also given a full explanation of the purpose of the study and a promise- which was fulfilled- of having the project share the research findings (Geismar 1973).

In an Australian study comprising samples of native families and immigrant families of European and Middle Eastern origin, motivation for participation, in addition to payment, included endorsement of the study by local public school systems and ethnic associations (Geismar and Geismar, 1979).

A different situation presented itself when interviewing (as part of a validity study) professional social workers for information about the social functioning of their clients (Geismar, Lagay, Niv, and Landau, 1986). The workers who shared an interest in the study results had no problem completing an 18 page checklist of scale items (Geismar and Camasso, 1993, pp.163–181). This contrasted favorably with the reaction of social work clients who responded positively to open-ended questions but often resisted interviews framed in a closed-ended format, which they found tiring.

Topic sensitivity is, of course, a major concern to all scale developers. A decision, whether given topics considered sensitive can be included, must be made in the pilot phase and pretest of the study. Yet a finding of subject suitability based on valid responses



from a one test population may not be applicable to a similar group of respondents, for reasons that are hard to predict. As an example we cite the differential response of sample populations to questions, under the heading of marital relationship, about the nature of sexual relations between partners (Geismar and Camasso, 1993, p.55). In our studies in the United States, Canada, Australia, and Sweden, we encountered occasionally resistance on the part of a respondent but not of a scope where we had to delete the subject entirely. The exception was a sample sub-population in an Australian study composed of Turkish immigrants. Vocal objections were raised by the translators and interviewers in this study who considered the inclusion of sexual content to be culturally taboo. We had to accept their judgment because lack of familiarity with language and culture left us no alternative but to omit the questionnaire segment to which objections had been raised.

#### Changes in Research Technology and the Quest for More Efficient Data Processing

During the mid-1950s when we first aggregated field data on the social functioning of multi-problem families we relied heavily on non-parametric statistical analysis, mindful of small samples and skewed distributions but also of computational simplicity inherent in the technique. Computers were only beginning to be available to social science data processing, but there were few programs that met our need. Scalogramming (Guttman scaling), then one of the more sophisticated methods of scale analysis had to be done manually (Geismar, La Sorte, and Ayres, 1962). We resorted to multi-dimensional scaling in the form of factor analysis when appropriate software was developed for mainframe computers in the late 1960s (Geismar, 1973). More recently the technique for testing reliability and validity received a boost with the use of structural equation modeling (Camasso and Geismar, 1992; Geismar and Camasso, 1993).

The point being made here is that changing research technology affords an opportunity to strengthen measurement by improved ways of handling data. Improvement may be in the nature of extracting information not heretofore available or processing the data more efficiently.

Access to more advanced research technology is by itself not a sufficient reason for modifying procedures. The technology has to be fully geared to the methods, goals, and resources of the study. Factor analysis, for example, always impressed us as an effective mode of data reduction. But were hundreds of hours of sitting behind calculators a desirable investment in manpower given the uncertainty of outcomes with varied population samples? Computer software once available made the testing of different factor analytic models a reasonable investment in time and expenditure.

Some changes in research procedure may be entirely unconnected to new research technology but motivated by the discovery, evolving from experimentation, that there are more efficient ways of processing data. For instance the most decisive modification in the use of the Family Functioning Scale occurred as a result of a search for a simplified way of coding narrative information. A twenty page rating guideline, which specified criteria for each main and sub-category, was replaced by a two page form whose criteria were at a higher level of abstraction. The simplified procedure resulted in a saving of time without sacrificing statistical accuracy. The revision eliminated some details of information from the routine statistical analysis without loss of reliability, but this information remained stored in the program for special focus studies and/or utilization by practitioners in the service program.

### Comment

This brief case history of the Family Functioning Scale was used as a basis for illustrating some of the common situations in which contextual factors help shape the nature of social and behavioral measurement. What should have become clear is that social scales and related instruments are not tools for all seasons but measures whose utility is constantly being challenged by changes in attitudes, demographic conditions, and social science know-how.

These observations have implications for researchers and students planning to make use of existing instruments of measurement, especially those that have been in circulation a number of years.

Instrument use, as stated earlier requires first and foremost an answer to the questions of reliability and validity. Researchers

tend to express a natural concern for the integrity of the measure and show a justified reluctance to subject it to modifications without further testing. This essay puts on the agenda the more complex issue of instrument relevance in the light of changing conditions. Typically, a research instrument is a one-shot endeavor that gets no further attention from its creators except in cases where they are engaged in longitudinal research or replication studies. In the case study reported here we had a situation where the principal investigators and associates opted to work on instrument development over a period of time.

Extended use of the instrument yielded information that transcends the conventional techniques for testing instrument adequacy focused in time and space. Data that accumulated over three and a half decades yielded a wealth of observations on the impact of social and other environmental factors upon the use of the Family Functioning Scale. These observations, hopefully, will stimulate similar studies in relation to other research tools.

The lack of opportunities for accessing information on extended use of measurement tools is largely due to a preoccupation with working on subjects that are new and original. Replication studies, which are standard procedures in the sciences, are neglected in social work and the social and behavioral sciences in general, where there is no industrial and medical infrastructure that relies on replication for boosting scientific rigor.

In an applied field like social work professional credibility hinges to no small extent on the prevalence of empirical practice research. The near absence of ongoing and coordinated efforts at instrument development presents a special barrier to such research. Efforts to overcome this deficit need to be guided by the recognition that in a clearly delineated field research effort in general and instrument development in particular are not solitary enterprises. Practice problems are shared by large segments of the profession and so should be endeavors aimed at scientific knowledge building.

Such sharing can take many forms, but at the most basic level efforts to assess practice scientifically should be undertaken on a cooperative basis, permitting a comprehensive process of instrument construction involving groups of social agencies and academic institutions on an ongoing basis. Ideally a central research

body, university or agency based, could serve as an integrator of such efforts. Cooperation in this realm would lead to a pooling of data, enlargement of samples, diversification of research methodology, and ultimately to the attainment of results that will strengthen the empirical knowledge base of the profession

### References

- Camasso, M. C. and Geismar. (1992). A multivariate approach to construct reliability and validity assessment: The case of family functioning. *Social Work Research and Abstracts*, 28, 16–26
- Geismar, L. L. (1971). *Family and Community Functioning*. Metuchen, NJ: The Scarecrow Press, Inc. 1971.
- Geismar, L. L. (1973). *555 Families. A Social-Psychological Study of Young Families in Transition*. New Brunswick, NJ: Transaction Books.
- Geismar, L. L. and Ayres, B. (1959). A method for evaluating the social functioning of families under treatment. *Social Work*, 4, 102–108.
- Geismar, L. L. and Ayres, B. (1960). *Measuring Family Functioning*. St. Paul, MN: Family Centered Project.
- Geismar, L. L. and Camasso, M. C. (1993). *The Family Functioning Scale. A Guide to Research and Practice*. New York: Springer Publishing Co.
- Geismar, L. L. and Geismar, S. (1979). *Families in an Urban Mold*. New York: Pergamon Press.
- Geismar, L. L. and La Sorte, M. (1963). Research interviewing with low-income families. *Social Work*, 1963, 8, 10–14.
- Geismar, L. L., La Sorte, M. and Ayres, B. (1962). Measuring family disorganization. *Marriage and Family Living*, 24, 51–56.
- Geismar, L. L., Lagay, B., Niv, S. and Landau R. (1986). Probing the validity of normative data. *Journal of Social Service Research*, 9, 69–76.
- Hunt, J. McV. and Kogan, L. (1950). *Measuring Results in Social Casework*. New York: Family Service Association of America.
- Overton, A. and Tinker, K. H. and Associates. (1957). *Casework Notebook*. St. Paul, MN: Family Centered Project.
- Task Force on Social Work Research (1991). *Building Social Work Knowledge for Effective Services and Policies*. Austin, TX.