Operant Subjectivity: The International Journal of Q Methodology
41 (2019): 93-109, DOI: 10.15133/j.os.2019.002

Operant Subjectivity
The International Journal of Q Methodology

From Concourse to Q Sample to Testing Theory

Steven R. Brown
Kent State University

Eric Baltrinic
The University of Alabama

Martin Jencius
Kent State University

Abstract: Misunderstandings have arisen about the nature and purpose of structured statement samples in Q methodology and their relation to concourses (the parent populations of statements) and to theory testing. Moreover, there is little discussion or understanding of proposition sets, or theoretical conditions under which participants are invited to perform Q sorts. A case study in program evaluation is presented as a way to illustrate procedures for collecting a concourse, converting its narratives into Q statements, and selecting a Q sample from the statement universe. An example is then presented of rationality in voting and how theoretical propositions are converted into conditions of instruction that provide facts bearing on theory.

Keywords: concourse, Q-sample structure, proposition sets, theory testing

In his *The Study of Behavior*, William Stephenson (1953) discussed at length the logic of the structuring of Q samples, and in later papers provided detailed consideration of concourse (Stephenson, 1978, 1986), from which Q samples are drawn. Only a few authors appear to use structured Q samples with any regularity, however, most composing Q sets in what seems to be an impressionistic manner, and fewer still employing propositional sets (Stephenson, 1953, pp. 42-46) or are even aware of their importance in the testing of theory, which is often associated with the Q-sample structure itself. It is the intent of this article to clarify these concepts — concourse, Q sample, and propositional sets — and their connection to the testing of theories about subjectivity.

A Case Study of Program Assessment

The Concourse

For almost two decades, secondary school teachers from around the world have been invited to participate in a program offered at Kent State University and other universities — the International Leaders in Education Program (ILEP), funded by the U.S. Department of State — that provides participants with courses, internships, and exposure to current educational technology. Upon return to their homelands, ILEP graduates have been asked to provide feedback concerning how they fared when they endeavored to implement concepts and practices to which they had been introduced while in the program. Specifically, they were asked to respond to the following questions:

Contact author: sbrown@kent.edu

1. What are some of the challenges you faced returning to your home country and trying to implement your ILEP plan?
2. If you overcame the challenges (mentioned in the previous question), how did you do that?
3. What successes have you seen as a teacher leader?
4. Have you received any recognition or awards for your teacher leadership since returning home? If yes, what recognition did you receive and when?
5. Did you have opportunities to share about American culture or other ILEPers' culture since your return? And if so, to whom, and how did you share?
6. After your return, did you have ongoing projects with Americans or other ILEP cultures? If so, please briefly explain.
7. In what ways did the technology training you received here impact your school/community?
8. As you reflect on your experience, what was the most impactful experience on you personally? In what way, and why?
9. Do you have any other comments or insights for us about the experience as an ILEP scholar and your impact on your school?

The narratives generated by the 48 alumni who responded over the years were dutifully recorded in an Excel file without regard to what might be done with this material from the standpoint of analysis (further details are reported by Jencius, Baltrinic, & Brown, 2017). This collection gradually amounted to more than 45,000 words, and it is this body of verbal roughage that constitutes a *concourse of subjective communicability* (Stephenson, 1978, 1986), which has been referred to as "perhaps the most important concept in the methodology" (Stephenson, 1980a, p. 99), for it is in concourse that life is primarily lived — where affections are expressed, plans are discussed, disagreements are resolved, and solutions to problems are conceived and explored — and it is concourse that provides the basis for a science of subjectivity. Not all of the ILEP comments were of a subjective nature, of course; many statements of fact were also in the mix. The following is an illustrative comment from one participant when responding to question no. 1 above:

> When I returned to my home country, I was reinstated as the Head of School of the senior section. I had the responsibility of taking care of the academic issues as well as the social issues, even teacher recruitment. I was lucky enough to work with a very broad-minded Principal who gave me the opportunity to share my newly acquired ideas. While trying to implement the new ideas I felt that I needed more knowledge about modern technology, such as the computer.

Being reinstated to a position as head of school is a statement of fact that is subject to testability as to its truth or falsity, but nestled within the above testimony are also to be found statements of opinion and feeling, *viz.*:

> I have been lucky to be supported by administrators who have given me the opportunity to share my new ideas.

> I have sometimes felt that I need more knowledge about the computer and other modern technologies.

With only slight editing, the above two statements remain faithful to the participant's original utterances and are also rendered applicable (i.e., non-specific) for use by other participants when describing their own back-home experiences. This extraction of the subjective gold from the ore of primitive narratives resulted in a reduction to more than 550 opinion statements that covered a wide range of perspectives and experiences, examples of which appear in Table 1.

Q-Sample Structuring

Since concourses are in principle infinite in magnitude and consequently impossible to gather in their entirety in most instances, pragmatics suggest the gathering of a reasonably large number of elements (such as the $N = 562$ in this case) to serve as a proxy for the concourse and from which the Q sample is to be drawn.[1] The common problem facing researchers involved in Q-methodology studies is how to reduce the voluminous concourse to a Q sample that is small enough for practical purposes and sufficiently diverse as to approximate the diversity of the concourse — i.e., to achieve *representativeness* of the stimulus domain (Brunswik, 1947; see also Brown & Ungs, 1970; Stephenson, 1953, p. 77) — and it is this problem for which the strategy of Q-sample structuring was designed to provide at least a partial solution. Note that reducing a concourse to a Q sample of smaller magnitude does not imply *reductionism* in the sense of explaining a more complex phenomenon in terms of its constituent parts (e.g., molecules in terms of atoms). A Q sample is merely a reduction in size of a larger concourse in the same way that a sample in survey research is a reduction of a population of persons.

In earlier days, Q samples were sometimes composed by selecting statements randomly from what then was referred to as the statement universe, a strategy that Hilden (1958) followed and that even Stephenson (1950, p. 32) pursued at one point; however, it was quickly recognized that a universe that was skewed would produce a Q sample similarly skewed. In his study of Jung's introversion/extraversion typology, for instance, Stephenson (1953) said:

> Jung provided many more statements descriptive of introversion than of extroversion, and more *thinking* and *feeling* functions than of the others. Any randomly chosen sample would therefore be biased by those predilections of Jung, and a balanced block design merely sorts the issues out in an explicit manner, so that all such possibilities will be taken care of equally…. (pp. 70-71)

One of the purposes of Q-sample design, therefore, is to counter biases and the selectivity of the proxy statements (which, for notational purposes, might be referred to as the *second-order concourse*); i.e., not to correct for biases in the broader concourse itself (the *first-order concourse*), which is infinite and exhaustive, but in those

[1] Acquiring a surrogate concourse that is broadly representative requires undertaking interviews with a diverse set of respondents who are pertinent to the concourse and structuring the interviews so as to cover a wide range of salient issues. Alternatively, drawing statements from an already-existing source has often been relied upon when the source can reasonably be assumed to have exercised comparable care. Newman's (2005) study of doctor-assisted suicide is illustrative of the latter strategy, relying on statements taken from studies by Fins and Bacchetta (1994, 1995), who in turn had broadly surveyed major statements on suicide from Kant to the present day. Additional considerations have recently been discussed by Gen (2020), Kenward (2019) and Sæbjørnsen et al. (2016).

statements that have been obtained from interviews, reviews of the professional literature, and other sources and therefore become candidates for inclusion in the Q sample. It is in selecting the second-order concourse that biases can enter and it is this that Q-sample structuring can partially neutralize.

Table 1: Mining Q Statements from Raw Verbiage

In response to question 1. Challenges Faced upon Return Home

From narrative text…	→	… to concourse statement
The red-tape bureaucracy has been the major stumbling block on the way of implementing projects in Morocco. Providing papers and documents from the sponsoring party, papers of and permission for any visiting guest and so on are real challenges and hurdles.		Bureaucratic red tape has been a stumbling block to implementing projects. Providing documents, getting permissions, etc., have been real challenges.
Shifting my fellow teachers' teaching paradigm, encouraging teachers to adopt information technologies, and promoting the use of technology in learning and teaching.		Moving old paradigms and getting others to adopt information technologies have been a problem.
When I returned to my home country, I was reinstated as the Head of School of the senior section. I had the responsibility of taking care of the academic issues as well as the social issues, even teacher recruitment. I was lucky enough to work with a very broad-minded Principal who gave me the opportunity to share my newly acquired ideas. While trying to implement the new ideas I felt that I needed more knowledge about modern technology, such as the computer.		**I have been lucky to be supported by administrators who have given me the opportunity to share my new ideas.** I have sometimes felt that I need more knowledge about the computer and other modern technologies.
I talked to my students and colleagues about different exchange programs that the American Embassy of Beirut is offering. I organized and did 2 workshops with the British Council about leadership in education…. Another problem is that some of the people I talked to were misled by a different image of the American view to Lebanese people than the one I encountered during my stay at the States. A third problem was that students tend to think that there is no more beautiful tomorrow or that there is nothing outside the small community that we are living in, and hence I talked to them about another country, in short, I gave them hope.		Many of my friends and associates seem to have been misled into a different image of Americans than the one that I encountered during my stay. **I have been able to give students hope in a more beautiful tomorrow, in something outside the small community in which they are living.**

Note: Statements in bold were eventually incorporated into the Q sample. A similar approach is taken by Paige and Morin (2015, p. 106).

Another purpose of Q-sample design is to *model*, or *simulate*, the statement universe; i.e., to provide a formal representation of how the concourse might appear from a theoretical standpoint, given that the statement universe is in most cases infinite and inexhaustible, hence of unknown and unknowable boundaries. Referring again to Stephenson's (1939, 1953) study of Jung's typology, the assignment of statements according to their introversion or extraversion (and then again in terms of whether they represented thinking, feeling, sensing, and intuiting functions as well as unconscious vs. conscious awareness) was to establish a starting point — an initial step in a more complex dependency analysis — and not to propose a measurement device in the sense of a scale of known meanings and dimensions.

It was this feature of Stephenson's thinking that was perplexing to Block and others:

> To most psychologists, it seems that the statements are only obscurely related, if at all, to the categories they nominally represent. On what basis, for example, can the items, "Impulsive and unrestrained," "A 'prophet,'" "Is underestimated and misunderstood," all be assigned to the same category; *introversion-unconscious-intuition*? What reasoning underlies the assignment of the items "Ponderous" and "Unreasonable" to the *extroversion-conscious-feeling* cell? Clearly, the reproducibility by other psychologists of the classification of items would be extremely low — a devastating defect in the approach. (Block, 1961, pp. 50-51)

This concern for accuracy, standardization, and certainty of meaning was at the heart of the matter and continued to bother Block (2008) in a revision of his earlier text a half century later: "Especially, Stephenson considered it 'a mistake to regard a sample as a standardized set or *test* of statements...' ([Stephenson], 1953, p. 77) — a view contrary to the present orientation" (Block, 2008, p. 32). As intimated previously, however, it was never a goal to achieve reproducibility, certainty, or standardization; rather, to use the structural logic of Fisher's (1935) principles of experimental design to achieve stimulus representativeness (see also Brown, 1970, 1971; McKeown & Thomas, 2013).

This worry over linguistic clarity and precision was not Block's alone. Sundland (1962), for instance, expressed disquiet about possible synonymous meanings among various statements, which, he surmised, could artificially elevate correlations, but as Stephenson (1963, p. 269) pointed out, language-in-use takes place "on concrete 'fields' of action" that override the logic of dictionary definitions — hence, a child may be *shy*, a blushing bride *bashful*, and a spinster *modest*, each term acquiring connotative and non-synonymous meaning in a specific context — a principle worth keeping in mind when it comes to selecting statements based on concepts comprising a Q-sample structure. Wittenborn (1961), in league with Block, noted that statements in Stephenson's Q samples "appear to have been somewhat informally assembled, and as a consequence, the analyses performed on the sorts…have an uncertain meaning" (p. 138). Wittenborn's remedy was to R-factor analyze the statements first — presumably to determine their true underlying meanings *qua* R-factor associations — and then to structure the Q sample on the basis of the R factors. Neff and Cohen (1967) similarly rejected Stephenson's apparent desire to assign statements to a Q-sample structure by "investigator's fiat" and sought to implement Wittenborn's suggestion — via the mechanics of variance analysis, however, rather than R factor analysis — and provided a scheme for determining whether any given Q sorter performed in a way consistent with the theory built into the structured sample. Finally, Kirschbaum, Barnett, and Cross (2019) have sought to assure greater clarity by inviting subject-matter experts to

cast votes (using a Likert scale via an online Delphi process) for those items from a concourse that they regarded as most important to include in the Q sample, which consequently ends up being of doubtful representativeness.

If Block et al. have conveyed a false narrative by overemphasizing systematics, then Watts and Stenner (2012) appear to be at the opposite end of the same narrative in rejecting the structuring of Q samples as representing "a clear sense of system and rigidity…that often appeals to researchers approaching Q from a quantitative background" (p. 59). By failing to see the value of Q-sample design as a means other than for quantification, however, Watts and Stenner end up throwing the baby out with the bath water and then find it necessary to come up with an alternative scheme for achieving representativeness. Their recommended strategy is to select elements from the concourse *as a whole* in such a way as to represent the statement population "smoothly and effectively without overlap, unnecessary repetition or redundancy" (p. 59), a worthy goal for which, however, they provide no guidelines and that consequently brings us back to the investigator-centric position that Q-sample structuring was designed in part to avoid

.

The ILEP Q Sample as Illustrative

The structuring of the ILEP Q sample provides an illustration. The statements of the concourse (examples of which are shown in Table 1) generally revolve around the issue of educational leadership, as did the ILEP program, and so the initial impulse was to consider structuring the Q sample in terms of the concepts and theories of leadership, and this eventually led to a consideration of the well-known conceptual framework found in a leading text by Kouzes and Posner (2018), *The Leadership Challenge: How to Get Extraordinary Things Done in Organizations*, the main features of which are summarized in Table 2. Kouzes and Posner propose five leadership practices, each of which being divided into two leading commitments (in italics), as follows:

Table 2: Q-Sample Structure (from Kouzes and Posner)

Practices	**Commitments**	
1. Model the Way	Clarify values	Set the example
2. Inspire a Shared Vision	Envision the future	Enlist others
3. Challenge the Process	Search for opportunities	Experiment and take risks
4. Enable Others to Act	Foster collaboration	Strengthen others
5. Encourage the Heart	Recognize contributions	Celebrate values and victories

(1) Model the Way: *Clarify values* by finding your own voice and affirming shared values, thereby promoting loyalty and teamwork. *Set the example* by personifying shared values and teaching others to model the team's values.

(2) Inspire a Shared Vision: *Envision the future* by imagining the team's possibilities — reflecting on the past, attending to the present, and imagining the future. *Enlist others* by aligning dreams and goals to shared ideals related to a common vision of the future.

(3) Challenge the Process: *Search for opportunities* by being innovative and creative and by seizing the initiative in developing solutions. *Experiment and take risks* by learning from successes and failures and by facilitating a climate that can generate small wins and larger results.

(4) Enable Others to Act: *Foster collaboration* by creating a climate of trust and facilitating relationships in joint projects. *Strengthen others* by enhancing self-determination, competence, and confidence; i.e., by relinquishing power.

(5) Encourage the Heart: *Recognize contributions* by expecting the best, encouraging feedback, and personalizing recognition and appreciation. *Celebrate values and victories* by creating a spirit of community through showcasing effort and applauding achievement.

Given the 5 practices with 2 commitments each resulted in 5×2 = 10 cells. Having adopted Kouzes and Posner's comprehensive framework, each of the more than 500 statements contained in the concourse were first placed in one of the five main categories; i.e., were placed in that category (Modeling the Way, Inspiring a Shared Vision, etc.) that seemed the best fit. For example, the following two statements — (*i*) "I have become more tolerant and respectful of others' point of view" and (*ii*) "I am more willing to try new approaches to teaching" — were initially placed in the category "Modeling the Way." The distribution of statements was naturally uneven, with 189 statements in Model the Way, 111 in Inspire a Shared Vision, and so forth. The statements in each category were then divided between the two best-fitting subcategories; e.g., statements (*i*) and (*ii*) above were placed in the subcategories "Clarify values" and "Set the example," respectively.

The final step in Q-sample composition involves selecting statements from each of the 10 subcategories shown in Table 2, and this is accomplished in accordance with the *principle of heterogeneity*.[2] Of the 91 statements in the category Challenge the Process, for instance, 31 were in the subcategory "Experiment and take risks," and of these, the following *m*=5 were selected that were as different as possible:

8. I have taken the initiative and created new learning opportunities for my students.
11. More than before, I now tend to suggest creative solutions.
29. Implementing new technology in the classroom requires overcoming multiple challenges at the individual and system levels.
35. I teach about other cultures in an effort to encourage educational paradigm shifts.
46. I have learned to be a risk-taker — to move out of my comfort zone and be willing to "jump into the fire." This is how you find out whether you can make mistakes and know how to fix them.

[2] Statements were initially placed in each of the 10 subcategories on the basis of their similarity (i.e., according to the *principle of homogeneity*): all in the first cell in Table 2 having to do with clarifying values, all in the second cell with setting an example, etc. Within-cell sampling on the basis of heterogeneity serves to counter the systematics of the Q-sample structure, thereby broadening diversity and more closely approximating representativeness.

m=5 statements were likewise selected in the same way from each of the other cells, resulting in a Q sample of size *N*=50 (i.e., 5 statements as replications from each of the 10 cells), with all statements then being randomly numbered, which further helps disguise the sample structure from the participants.

Factor-Analytic Results

Were a "test of the statements" at issue, as Block seems to have assumed, then a variance analysis of the statement scores might be applicable and we would be seeking to determine (for instance) whether "envisioning the future" was more important than, say, "fostering collaboration" (from Table 2), or "sharing vision" significantly greater than "encouraging the heart." But these categories from Kouzes and Posner and the assignment of statements to them are matters of *logic and reasoning on the part of the investigator* — e.g., that this particular statement is more compatible with category X than category Y — whereas the scores arising from the Q sorting (and to which variance analysis would be applied) are matters of *feeling and self-reference on the part of the participants*, who have no knowledge of Kouzes and Posner and are not involved in logically categorizing statements, but are busy representing their views concerning their experiences back in their home countries. Any analysis from the standpoint of a science of subjectivity must therefore take as central this self-reference, which is absent from any variance analysis of constructed effects (Brown, 1999). As Stephenson (1993-1994) concluded:

> There is therefore an overriding consideration for sweeping variance analysis aside, and to accept, instead, the more profound method of factor analysis. Factor theory, in the Q-form (but not R), maintains self-reference at the center of all else. Its "effects" are self-references. (p. 13)

Briefly, what the factor analysis shows in this case is summarized in the following statements and their factor scores (for factors X, Y, and Z, respectively):

+4	0	–2	[a]	I have tried to incorporate new technology into my work and have urged colleagues to become involved in our growing ICT [information and communications technology] culture.
+2	**+4**	+2	[b]	ILEP empowered me to dream bigger and opened me to the reality that, as an educator, I have a lot to learn and a lot to share, too.
0	–4	**+4**	[c]	The ILEP experience led to my being recognized with awards, career advancements, etc.

The first of the three factors was comprised of the *Technologists* who, as shown in statement [a], were keen to take their new knowledge of computers and to incorporate it into their work and to urge others to do the same. It can no doubt be argued that elements of the Kouzes and Posner model can be detected in factor X — of modeling the way, of enlisting others, fostering collaboration, etc. — but these effects were not in the foreground. What was at the focus of this factor's attention was enthusiasm for newly-acquired knowledge of technology and championing its usage in an emerging technological culture. As to factor Y, the *Grateful*, these participants were appreciative of the ILEP program (witness statement [b] above) for equipping them to be better

leaders in their organizations, yet gratitude was not *in the Q sample* as such: it was in the participants' grateful feelings *imposed on the statements* by these thankful Q sorters who now felt that they could act with a more empowered sense of effectiveness in their professional lives. Finally, factor Z, the *Accomplished*: in this factor array, these participants, as in statement [c], were displaying their successes, thereby indicating that they required no further assistance from the ILEP staff, a question that had been a major impetus for the program assessment.

In Conclusion

The principles and procedures sketched out above can be summarized in terms of the following graphics. Assume that the *abcd* rectangle in Figure 1 represents the entirety of all that can be said about some topic; i.e., represents the concourse. And assume further that we elect to structure the concourse in terms of a theory or conceptual framework that can be represented by a 3×4 factorial design, where one effect contains three levels (ABC) and the other four levels (WXYZ), resulting in 12 cells into which each of the 250 elements can be placed. The numbers in the cells indicate the hypothetical number of statements placed in each category (e.g., 15 statements of the AW variety, 20 of the AX variety, etc.). The fit need not be perfect in every case and disagreements about some of the 250 placements can be expected, but it is important to grasp that there is to be no test of these item assignments. It is certainly unnecessary to obtain agreement from a committee, as assumed by Block (1961), Kirschbaum et al. (2019), and others, although comparing notes with an associate can be beneficial. Note that the empty space outside the (ABC)×(WXYZ) interior rectangle represents statements in the overall concourse that were unavoidably omitted in the 250 statements that were gathered and from which the Q sample will be selected. Concourses are ordinarily infinite in size and so incapable of being included in their entirety. Omissions are inevitable.

a *b*

	A	B	C
W	15	20	25
X	20	25	30
Y	25	15	20
Z	15	30	10

c *d*

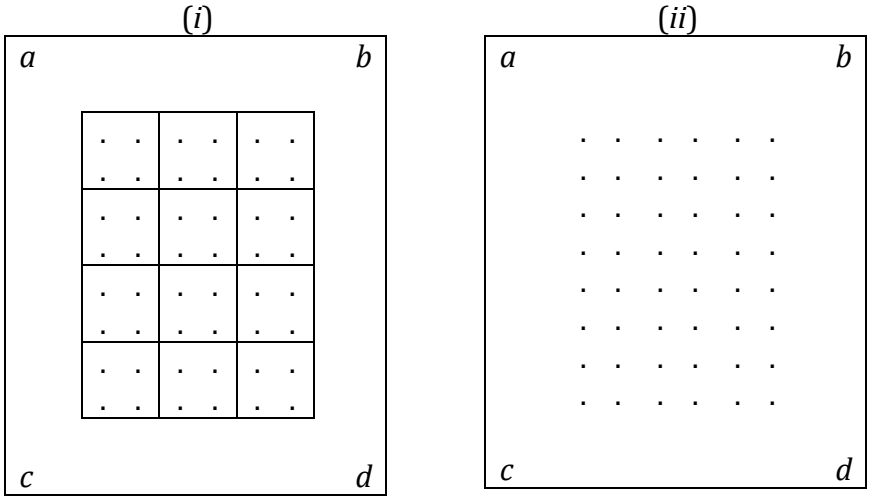
Fig 1. Concourse (N = 250).

Parenthetically, theories or conceptual frameworks chosen as the basis for Q-sample design should be of a broad nature suitable for comprehensively covering the concourse. Jung's (1923) introversion-extraversion typology, Erikson's (1963) stages of psychosocial development, Holland's (1992) vocational types, Kouzes and Posner's (2018) leadership framework, and the like are of this scope and therefore provide a place for almost all statements in their respective domains. But even in situations in which investigators claim that the literature provides no theory in which to couch their study, more general conceptual frameworks are usually available.[3]

[3] Of general utility, for instance, is Lasswell's (1971, pp. 24-25) conceptual framework for *perspectives*, which include *demands* (for action of some kind), *identifications* (of the self and not-self with persons, institutions, concepts, etc.), and *expectations* (allusions, often causal, to past, present, and future events). A Q sample about the threats of avian influenza (AI) included statements such as the following (Brown & Wattanakul, 2008): "Laboratories using new and improved tests for AI must implement appropriate quality management systems that provide veterinary authorities with confidence in the validity of test results" (*demand*), "Stamping out policies have led to very high costs and economic

Given the 3×4 = 12 cells, a decision is made to select *m*=4 statements from each (for a Q sample of size *N*=48). Turning to the 15 elements in the AW cell, 4 of the 15 available elements are selected according to the principle of heterogeneity; i.e., *m*=4 are selected with the goal in mind to create diversity from among these 15 statements that are alike in kind in that they are all of the AW variety. This is pictured in *Figure 2(i)* by the four dots within the cell that are as far apart as possible rather than being bunched together. The same logic is applied to each of the 12 cells, resulting in 48 dots, with each column, each row, and each cell of the design equally represented. Removal of the cell boundaries in Figure 2(*ii*) shows the Q sample spread out across the concourse and as being representative of it. It is for this reason, primarily, that the participant, when operating with the Q sort, typically detects no structure and experiences the statements as reasonably comprehensive of the substantive terrain.

Fig 2. Q sample (N = 48).



The principles advanced by Block (1961), Wittenborn (1961), and others mentioned above — and even those recommended by Watts and Stenner (2012) — are not apt to achieve representativeness as assuredly as the procedure outlined above. Watts and Stenner, for example, wish to place all 250 of the Figure 1 items before them and then to select a diverse set of, say, 48 statements from the 250, but this is an enormous task (although perhaps not recognized as such) for which the human mind is ill equipped,[4] and leaves the observer vulnerable to violating Feynman's (1985) first principle, "… that you must not fool yourself — and you are the easiest person to fool" (p. 343). It is more manageable to examine the 15 statements from cell AW above and then to select 4 that are diverse (e.g., dissimilar in language and expression so as to diminish resemblance), and then move on to the next cell and do the same. This morselizes the task and renders its accomplishment more likely.

---

losses for the affected nations, stakeholders, and, ultimately, consumers" (*identification*), and "Given today's frequent international travel and cross-border flow of goods, once a pandemic strain does emerge, it will be impossible to prevent it from spreading globally" (*expectation*). This framework can be employed in virtually any setting.

[4] Even for the most populous of the cells in Figure 1 (BZ or CX, with *m*=30 items), from which there are ${}_{30}C_4$ = 27,405 ways to select *m*=4 items, there more than 296 quindecillion times as many ways to select 48 items out of all 250 (1 quindecillion equals 1 followed by 48 zeros), which leaves practically infinite room in which to err in judgment.

Finally, we are reminded of Wittgenstein's (1922, prop. 6.342) warning that theories are like meshes of net imposed upon a picture, much as portrayed in Figures 1 and 2(*i*) above, such that describing what we see square by square tells us more about the net than about the picture beneath it. This is what variance analysis does: it describes a person's Q sort (or a factor array) in terms of the Q-sample structure rather than in terms that are intrinsic to the person's response. Q factor analysis, by way of contrast, deals directly with the picture in Figure 2(*ii*) and reveals what is *operant* within it (Stephenson, 1977) without regard to the constructed effects built into the Q sample a priori.

Structured Q Samples, Propositional Sets, and Theory Testing

Criticism of structured Q samples has generally been based on imperfect understandings of Stephenson's intent. After referring to earlier practices, for instance, Block (2008), comments as follows: "A third, ambitious Stephenson approach aspired to theory testing through an analysis of variance design said to be embodied by a 'structured' set of Q statements. This aspiration proved to be doomed by its failure of credible implementation" (p. 31n). This assumption — that explanations are contained in the statements or testing device — was protopostulatory to critics such as Wittenborn and others reviewed earlier and yet remains embedded in current thinking, even among devotees of Q methodology; e.g., "Structured Q sets can be used to test the appropriateness of theory in conceptualizing subjective views" (Baker, McHugh, & Mason, 2017, p. 173). However, theory testing via statement design was never an intention as Stephenson (1953, 1993-1994) made clear throughout his career.

The motivation to carry out a study is usually prompted by a theory of some kind, but if theory is to be used to help fashion the Q sample and then dismissed thereafter, then how and at what point does testing enter the picture? Stephenson was explicit about this — detailed illustrations can be found scattered throughout *The Study of Behavior* (Stephenson, 1953, consult *Propositions* in Index) as well as other locations (e.g., Stephenson, 1976; 1980b, pp. 10-11, 16-18, 24-28; 1992) — but his views on these matters have for the most part been left unattended. The Q sample can be implicated in the testing situation, of course, but there is a change in direction at this point that involves turning from the measuring device itself to an examination of actual behavior under controlled conditions, as illustrated in the example below.

Before turning to Stephenson's views about theory testing, however, it is worth asking: From where did this impression arise that theory proceeds in terms of the Q-sample structure? Part of this is due to the history of testing theory in the human sciences in general, where testing has traditionally involved the use of measuring instruments such as rating scales: it was easy to assume that Q sorts were like rating scales, an assumption in wide circulation even now (witness the occasional article comparing Q to other scale techniques [e.g., Serfass & Sherman, 2013]). Part of the blame can also be placed at the feet of Stephenson himself, whose *The Study of Behavior* devotes considerable space to Q-sample structuring and variance analysis. (He once remarked in a graduate seminar in the mid-1960s that he probably overemphasized variance analysis but felt that American psychology at the time seemed blissfully unaware of the contributions of R.A. Fisher.)

One of the greatest contributors to variance-analytic thinking in Q methodology was Kerlinger (1964), whose authoritative *Foundations of Behavioral Research* contained the first chapter-length treatment of Q methodology and was enormously popular and

widely adopted beyond the field of education. In it, Kerlinger stated, "… the basic rationale of *Q* as Stephenson sees it is that individuals sort the cards not so much to test the individuals as to test 'theories' that have been built into the cards" (p. 587). The fact that Kerlinger (1972) contributed a chapter to the Stephenson *Festschrift* no doubt added authority to his views. Possibly the strongest statement of Kerlinger's emphasis on variance analysis and the structuring of Q samples is in his little-known monograph *Q Methodology and the Testing of Theory* (Kerlinger, 1958), a 55-page step-by-step guide to structuring statements and analyzing them via variance and factor analyses. But perhaps the clearest testimony to Kerlinger's influence is to be found in the Q *Primer* (Brown, 1993), in which Kerlinger's view and that of one of his followers formed the basis for a separate Q factor, orthogonal to the factor containing Stephenson's position, about the nature of Q methodology. Kerlinger's was a different understanding of Q methodology.

Stephenson's view about testing, expressed most thoroughly in Chapter 2 of *The Study of Behavior*, was that it was to take place primarily in terms of Q sorts performed by *single cases* or small numbers of them under various *conditions of instruction* informed by *propositional sets*, which are assemblages of assertions related to a phenomenon (e.g., that an extravert's self and ideal self will be congruent, that an introvert's ideal self will be extraverted, that an introvert will have more insight into an extravert than vice versa, etc.). For his ideas about propositions, Stephenson relied heavily on the Vienna positivists, and also on Wittgenstein (1922), to the effect that propositions are assertions about the world that are then compared to facts (Schlick, 1935). That some people are extraverts and others introverts, for instance, is one such proposition, for which a Q sample structured in terms of Jung's typology would be used, but this is only a *universal*, or general, proposition (Kaufmann, 1944, p. 24), which is to be distinguished from a *singular testable* proposition, such as that "the introvert's ideals are extraverted." Evidence regarding the latter is obtained by inviting an introvert to provide a Q-sort description of how she might like for others to see her and then to see whether the facts agree with the proposition (for illustrations, consult Brown & Hendrick, 1971; Expositor, 1985; Stephenson, 1953, pp. 42-46, 157-158).

Many, like Kerlinger, erred by mistaking for a general theory what were only universal propositions — e.g., that some people are introverts and others extraverts, or that some are theorist types and others economic types from Spranger's theory, to use Kerlinger's (1964, pp. 587-589) own example — ignoring entirely the myriad singular propositions laid out by Jung and Spranger and in terms of which the types were to be understood and explained. It is often the general proposition that is built into Q samples or into devices like the Myers-Briggs Type Indicator, and usually on logical grounds — e.g., that statement *x* is introverted, that statement *y* is extraverted, etc. — but the testing of singular propositions is carried out by examining actual behavior; i.e., in observing how individuals behave under various conditions of instruction that serve to put the proposition to test.

A worked example is focused on the concept of the rational voter (for details, see Brown, 1993-1994), about which numerous propositions have been advanced — e.g., (a) that the rational voter is mindful of the gains to be had and the costs to be borne by voting one way or another and has a desire to go where rationality leads, (b) that the rational voter takes into account the relative benefits to be expected from political parties in competition, both before and after elections, and (c) that the rational voter implicitly compares the performance of the party in office to an ideal political party in

the same position. But do these propositions really describe the way in which citizens think when they make voting decisions?

In a case in point, a single voter was provided with a Q sample containing a variety of policies that a government might pursue and was invited to operate with them under a variety of performances — (1) What are your own policy preferences? (2) Which policies, if enacted, would benefit you most? (3) Which policies would cost you the most? (4) Which policies are the Republicans currently promoting in their campaign? (5) Which policies will the Republicans pursue, if elected, after the election? (6) and (7) the same for the Democrats? (8) What policies would be pursued by an ideal government? As with structured Q samples and P sets, proposition sets are designed to be of sufficient representativeness to render it likely that major factors will have the opportunity to be revealed. The resulting three-factor structure is in Table 3, based on a theoretical rotation that gave emphasis to the person's own preferences, at which point it was discovered, somewhat surprisingly, that benefits and costs were orthogonal to preferences.

Table 3: Operant Factors of Voting

Experimental Conditions	**A**	**B**	**C**
1. Personal preference	**X**		
2. Gains		**X**	
3. Costs			**X**
4. Republicans now	**–X**		**X**
5. Republicans after election	**–X**		**X**
6. Democrats if in office	**X**		**X**
7. Democrats after election	**X**		**X**
8. Ideal government	**X**		

The findings clearly show that considerations of benefits and costs do not enter functionally into this person's electoral calculus, which led to the replacement of the concept of *rationality* with the *induced* proposition (Stephenson, 1953, p. 46) of *rationale*: this voter has her own reasons for voting the way she does. She is aware of potential benefits but does not embrace these policies as a matter of personal preference. Knowledge of benefits are *mine*, she implicitly says — i.e., I am cognizant that certain policies would be to my benefit — but not *me* (James's Law). The person's preferences are also aligned with her view of an ideal government (Rogers's Law of self-ideal congruency), and the Democratic Party is recipient of this idealization. (On the laws of subjectivity, consult Brown, 2019.) The costs associated with various policies were implicated in both Republicans and Democrats (both before and after the election) as mixed cases in recognition that public policies cost money, and factor C would probably appear in anyone's and everyone's Q sorting. The question is whether the policies pursued are compatible or not with personal preferences, which explains the bipolarity of factor A: The Republican Party, according to this person, was apt to implement policies diametrically opposed to her own preferences.

Concluding Remarks

When in 1919 a group of astronomers sent a team to an island off the West Coast of Africa to observe the bending of starlight within the Sun's gravitational field (rendered visible by a solar eclipse), they were testing a proposition from Einstein's 1915 theory of general relativity that light is material (photons) and consequently affected by gravity, and that the fabric of space-time is warped by masses passing through it. The observers were not testing general relativity directly — rather, a singular testable proposition derived from it — and this is the way in which theories are tested, proposition by proposition. The more propositions affirmed (i.e., are failed to be rejected), the more confidence in the theory from which they were derived, and this holds true across science generally.

The mistake has been made historically to regard the structuring of a Q sample as the way to test a theory — e.g., by observing whether responses of specific categories of participants (such as businessmen vs. priests) comport with preferences for specific kinds of items (e.g., economic vs. religious statements) — rather than by examining behavior itself under controlled conditions, such as the conditions of instruction governing a Q-sort setting, where the instructions are surrogates for propositions arising from theory. This latter strategy necessitates single-case studies or studies of relatively small numbers of carefully selected participants rather than the Q sorting of large numbers of cases and constitutes an invitation to scientists interested in subjectivity to move in more closely so as to be able to examine behavior under improved circumstances.

It was Stephenson's good fortune to see at the outset how various principles and procedures in circulation in his time could all play a role in the development of Q methodology — of Brunswik's (1947) concept of ecological universes and of the need to embody them through representative design; of Fisher's (1935) variance designs for the theoretical structuring of Q samples as a means for achieving representativeness; of the theoretical clarifications about propositions by Wittgenstein (1922) and members of the Vienna Circle; of the recognition by Bridgman (1927) of the importance of operations and by Skinner (1938) of the possibility of extending this into the operantcy of behavior; of Kantor's (1959) and Dewey and Bentley's (1946) insistence on the concreteness of behavioral interactions, transactions, and segments; and of Spearman's (1927) factor theory (when inverted into its Q form) for providing the means for revealing the structures of these interactions and transactions. Due to space limitations, only a few of these concepts and principles have been discussed above and others have been added as Q methodology has evolved, and it is important for practitioners of Q methodology to be reminded of these foundational ideas from time to time.

References

Baker, R., McHugh, N., & Mason, H. (2017). Constructing statement sets for use in Q methodology studies. In J. Coast (Ed), *Qualitative methods for health economics* (pp. 163-174). London: Rowman & Littlefield International.

Block, J. (1961). *The Q-sort method in personality assessment and psychiatric research.* Springfield, IL: Charles C Thomas.

Block, J. (2008). *The Q-sort in character appraisal: Encoding subjective impressions of persons quantitatively*. Washington, DC: American Psychological Association.

Bridgman, P.W. (1927). *The logic of modern physics*. New York: Macmillan.

Brown, S.R. (1970). On the use of variance designs in Q methodology. *Psychological Record, 20*, 179-189.

Brown, S.R. (1971). Experimental designs and the structuring of theory. *Experimental Study of Politics, 1*(1), 1-41.

Brown, S.R. (1993). A primer on Q methodology. *Operant Subjectivity, 16*, 91-138.

Brown, S.R. (1993-94). The structure and form of subjectivity in political theory and behavior. *Operant Subjectivity, 17,* 30-48.

Brown, S.R. (1999). On the taking of averages: Variance and factor analyses compared. *Operant Subjectivity, 22*(3), 31-37.

Brown, S.R. (2019). Subjectivity in the human sciences. *Psychological Record*, *69*, 565-579.

Brown, S.R., & Hendrick, C. (1971). Introversion, extraversion, and social perception. *British Journal of Social and Clinical Psychology, 10*, 313-319.

Brown, S.R. & Ungs, T.D. (1970). Representativeness and the study of political behavior: An application of Q technique to reactions to the Kent State incident. *Social Science Quarterly, 51*, 514-526.

Brown, S.R., & Wattanakul, S. (2008, October). *Power, enlightenment, and threats to well-being: Perspectives and recommendations concerning avian influenza*. Presented at a meeting of the Policy Sciences Annual Institute, University of Colorado, Boulder, CO.

Brunswik, E. (1947). *Systematic and representative design of psychological experiments.* Berkeley and Los Angeles: University of California Press.

Dewey, J., & Bentley, A.F. (1946). Interaction and transaction. *Journal of Philosophy, 43*, 505-517.

Erikson, E.H. (1963). *Childhood and society* (2nd ed.). New York: Norton.

Expositor (1985). A note on "testing" psychological types. *Operant Subjectivity, 9,* 29-35.

Feynman, R.P. (1985). *Surely you're joking, Mr. Feynman! Adventures of a curious character*. New York: Norton.

Fins, J. J., & Bacchetta, M.D., (1994). The physician-assisted suicide debate: An annotated bibliography of representative articles. *Journal of Clinical Ethics, 5*, 329-340.

Fins, J. J. & Bacchetta, M. D. (1995). Framing the physician-assisted suicide and voluntary active euthanasia debate. *Journal of the American Geriatrics Society, 43*, 563-568.

Fisher, R.A. (1935). *The design of experiments*. Edinburgh, UK: Oliver & Boyd.

Gen, S. (2020). Developing the Q-sample and P-set in Q-methodology: An example from policy studies. *SAGE Research Methods Cases* [online]. London: Sage https://dx.doi.org/10.4135/9781529716092

Hilden, A.H. (1958). Q-sort correlation: Stability and random choice of statements. *Journal of Consulting Psychology, 22*, 45-50.

Holland, J.L. (1992). *Making vocational choices: A theory of vocational personalities and work environments* (2nd ed.). Odessa, FL: Psychological Assessment Resources.

Jencius, M., Baltrinic, E., & Brown, S.R. (2017, September 9). *The success and challenges of implementing teacher leadership after study abroad*. Read at the annual meeting of the International Society for the Scientific Study of Subjectivity, Glasgow, Scotland.

Jung, C.G. (1923). *Psychological types*. London: Routledge & Kegan Paul.

Kantor, J.R. (1959). *Interbehavioral psychology: A sample of scientific system construction* (2nd rev. ed.). Granville, OH: Principia Press.

Kaufmann, F. (1944). *Methodology of the social sciences*. London: Oxford University Press.

Kenward, L. (2019). A literature review to guide novice researchers using Q methodology in the development of a framework for concourse management. *Nurse Researcher, 27*(1), 17-21.

Kerlinger, F.N. (1958). *Q methodology and the testing of theory*. New York: School of Education, New York University. (mimeographed, 55 pp.)

Kerlinger, F.N. (1964). *Foundations of behavioral research: Educational and psychological inquiry*. New York: Holt, Rinehart, and Winston.

Kerlinger, F.N. (1972). Q methodology in behavioral research. In S.R. Brown & D.J. Brenner (Eds.), *Science, psychology, and communication* (pp. 3-38). New York: Teachers College Press.

Kirschbaum, M., Barnett, T., & Cross, M. (2019). Q sample construction: A novel approach incorporating a Delphi technique to explore opinions about codeine dependence. *BMC Medical Research Methodology, 19*(1), art. 101.

Kouzes, J.M., & Posner, B.Z. (2018). *The leadership challenge: How to get extraordinary things done in organizations* (6th ed.). San Francisco: Jossey-Bass.

Lasswell, H.D. (1971). *A pre-view of policy sciences*. New York: American Elsevier.

McKeown, B.F. & Thomas, D.B. (2013). *Q methodology* (Quantitative Applications in the Social Sciences series, Vol. 66, 2nd ed.). Thousand Oaks, CA: Sage.

Neff, W.S. & Cohen, J. (1967). A method for the analysis of the structure and internal consistency of Q-sort arrays. *Psychological Bulletin, 68*, 361-368.

Newman, T.D. (2005). *Links between ethics and public policy: A Q methodological study of physician assisted suicide and euthanasia*. Doctoral dissertation, Kent State University.

Paige, J.B., & Morin, K.H. (2015). Q-sample construction: A critical step for a Q-methodological study. *Western Journal of Nursing Research, 38*, 96-110.

Sæbjørnsen, E.N., Ellingsen, I.T., Good, J.M.M., & Ødegård, A. (2016). Combining a naturalistic and theoretical Q sample approach: An empirical research illustration. *Operant Subjectivity, 38*(2), 15-32.

Schlick, M. (1935). Facts and propositions. *Analysis, 2*(5), 65-70.

Serfass, D.G., & Sherman, R.A. (2013). A methodological note on ordered Q-sort ratings. *Journal of Research in Personality, 47,* 853-858.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century.

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. London: Macmillan.

Stephenson, W. (1939). Methodological consideration of Jung's typology. *Journal of Mental Science, 85*(355 [new series no. 319]), 185-205.

Stephenson, W. (1950). A statistical approach to typology: The study of trait-universes. *Journal of Clinical Psychology, 6,* 26-38.

Stephenson, W. (1953). *The study of behavior: Q-technique and its methodology*. Chicago: University of Chicago Press.

Stephenson, W. (1963). Independency and operationism in Q-sorting. *Psychological Record, 13*, 269-272.

Stephenson, W. (1976). Q-methodology: Conceptualization and measurement of operant effects of television viewing. *JCATS: Journal of the Centre for Advanced Television Studies* [London], *4*(1), 17-18.

Stephenson, W. (1977). Factors as operant subjectivity. *Operant Subjectivity, 1*, 3-16.

Stephenson, W. (1978). Concourse theory of communication. *Communication, 3*, 21-40.
Stephenson, W. (1980a). Michael Polanyi, science and belief. *Ethics in Science & Medicine, 7*, 97-110.
Stephenson, W. (1980b). Consciring: A general theory for subjective communicability. In D. Nimmo (Ed.), *Communication yearbook 4* (pp. 7-36). New Brunswick, NJ: Transaction Books.
Stephenson, W. (1986). Protoconcursus: The concourse theory of communication. *Operant Subjectivity, 9*, 37-58, 73-96.
Stephenson, W. (1992). Self in everyday life. *Operant Subjectivity, 15*(2), 29-55.
Stephenson, W. (1993-1994). Introduction to Q-methodology. *Operant Subjectivity, 17,* 1-13.
Sundland, D.M. (1962). The construction of Q-sorts: A criticism. *Psychological Review, 69,* 62-64.
Watts, S., & Stenner, P. (2012). *Doing Q methodological research: Theory, method and interpretation*. London: Sage.
Wittenborn, J.R. (1961). Contributions and current status of Q methodology. *Psychological Bulletin, 58,* 132-142.
Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. London: Kegan, Paul, Trench.