Reviewer’s report

Title: What Motivates Senior Clinicians To Teach Medical Students?

Version: 2  Date: 21 April 2005

Reviewer: Steven Brown

Reviewer’s report:

1 Is the question posed new and well defined?
The question posed is clear and important to assess the desire to teach medical students as a prelude to identifying and recruiting senior clinicians with this desire.

2 Appropriateness of method; replicability
Q methodology is quite appropriate for the question under consideration, but there is occasional awkwardness in expression that can be smoothed over through changes in word choice; there are also procedures that deviate from best practices but which might be judged good-enough in this context.
On p. 6 (under Methods/Study Group), the authors state that the study participants constituted a convenience sample, whereas on p. 7 (Choosing the Sample Population) it is said that all 75 available staff members were included, which means the study is based on the entire population (aside from those 26 staff members on leave). In most Q studies, the P set is selected by using experimental design principles (usually factorial), and something like the following might have been useful given the intent of this study:

Teaching Effectiveness
high mediocre poor
Specialty
surgery psychiatry radiology pathology

Given the $3 \times 4 = 12$ combinations, the authors might have selected $n=4$ surgeons with reputations for good teaching (perhaps based on student evaluations), $n=4$ with mediocre teaching credentials, and $n=4$ judged to be poor, and the same for the psychiatrists, radiologists, and pathologists (and any other categories). This would have provided the P set with great theoretical significance than characterizes a convenience sample or testing the entire population.

On p. 6, it is said that Q methodology was developed to bridge qualitative and quantitative research, but Q was invented (in 1935) well before this distinction, and an emerging view is that Q actually problematizes the division between qualitative and quantitative. (This was one of the conclusions expressed during a roundtable on Q-ing the Q: Views on the relationship between the quantitative and the qualitative in human inquiry, International Society for the Scientific Study of Subjectivity, September 23-25, University of Georgia, Athens.)

On p. 6, the Q sorting is characterized as the ranking of individual opinions about the statements, but the task is simply one of ranking the statements (not opinions about statements).
On p. 8, it is said that there are 69 Q statements, with equal numbers of positive and negative statements, which cannot be the case mathematically. As with the P set, the Q sample is conventionally structured in some kind of theoretical way, which is not done in this case. The following, for instance, were discussed on p. 4 and would have been apropos for this study:

Needs
physiological safety belonging esteem actualization
Goals
achievement pro-social
Rewards
extrinsic intrinsic

These 4x2x2 = 16 cells could have been replicated n=4 times for a total of N=64 statementse.g., n=4 physiological/achievement/extrinsic statements n=4 actualization/pro-social/intrinsic statementsbut now the Q sample would have a sharper theoretical edge.

As to the Q sample itself (Table 1), the language is somewhat stilted and in the future should be made to sound more natural. In natural language, participants wouldnt normally precede their assertions with the preface I dont teach over and over, as in statements 13, 15-17, 19-22, etc. This problem could have been solved in the present case by removing all of the statement prefaces, so that the Q sample would read something like the following:

1. spending time with students in small groups
2. lecturing to large groups of students
3. accepting the challenge of teaching students as effectively as possible
4. experiencing boredom while teaching
5. feeling a sense of duty to teach
6. setting a good example to students to become teachers
   Etc.

The participants could be asked to characterize their own experiences of teaching by ranking the items from most characteristic (+5) to most uncharacteristic (-5) of their experiences. Note also that statements that deal with a similar feature of the situation ought not be placed side by side, e.g., items 36 and 37. If items were numbered randomly, it would take care of this problem.

On p. 9, the authors state that there were no repeat measures taken, which I assume means that they did not ask their participants to perform the Q sort again at a later time. This would have constituted a replication, but hardly seems warranted. I would be inclined to delete this statement.

On pp. 9-11, the authors describe their data analysis and results. Inasmuch as this method is apt to be new to readers of BMC Medical Education, the authors might wish to consider referring readers to one or more of the following articles of the past decade:


On p. 10, it is said that Q methodology yields ipsative rather than normative data, but this distinction is not one that persons familiar with Q consider salient. That is, the concept of ipsative is generally found in the R-methodological literature that addresses Q-factor and Q-cluster analysis, but is not considered to be salient within the Q-methodological literature.

3 Are the data sound and well controlled?
The data themselves seem to have been carefully acquired under reasonable conditions. With regard to the data analysis, however, the authors do not indicate which of the two kinds of factor analysis that are built into the PQMethod program (principal components or centroid) that they used, nor the kind of rotation (probably varimax) was that employed. This would not matter to me so much, but other readers might be curious.
The fact that there are so few statements that distinguish factors 2 and 3 (Tables 3 and 4) makes me wonder whether these factors are highly correlated, either with one another or with other factors. The authors should probably check this.

4 Reporting and data disposition
I see no problems in these regard. The authors have included the Q sample (Table 1), which would be the main information that other scholars might want to have for purposes of replication or verification.

5 Adequacy of discussion and conclusions
The discussion section is quite spartan, but this may be the norm in medical writing. Most Q studies go into greater detail in the description of factors. In addition, I think that the authors have been too generous with their factor 1, which incorporates 90% of their participants. Factor 1 claims to enjoy teaching, and I would expect this view to predominate, even among many individuals who deep down do not enjoy teaching. Not many would be able to acknowledge otherwise; i.e., the demand characteristics of this experiment invite a bias of this sort. The authors briefly mention the so-called social desirability bias (p. 13), which, like the ipsative issue, has little credibility in Q studies generally, but they dont really take this issue and elaborate on its import in the context of this particular study. I dont think that this necessarily sinks the project, but the findings need to be qualified. That is, these factor results can be considered to constitute phase-1, in which participant responses are initially accepted at face value, with more careful discernment (if need be) to follow in subsequent studies.
The authors mention so-called limitations of sampling, reliability, and the quasi-normal shape of the Q-sort distribution (p. 13), but give no hint of the solid responses to these criticisms that exist in the Q-methodology literature.

6 Accuracy of title and abstract
Both title and abstract seem to imply that the Q sort captures motivation. For reasons given above, the authors might want to be more tentative about this. A more accurate title might be something like Senior Clinician Perspectives on Motivations for Teaching Medical Students.

7 Acceptability of writing
The writing is clear and orderly. Reference 10 (p. 17) does not give full bibliographic details (especially the authors name) of this conference presentation.

**Recommendations for Revision**

Discretionary. Much of what has been said or implied above may be considered discretionary. The best practices recommendations about Q-sample and P-set design are for the authors consideration in their next Q study. If Q methodology is not well known to readers of the journal, then adding some of the references noted above might be useful to readers and also serve to help legitimate the current study. I would also leave it to the authors as to whether they wish to adhere to the notion that Q yields ipsative data.

Minor essential. Some of the defects mentioned above are only a matter of wording and can easily be rectified e.g., the difference between convenience sampling vs. querying the entire population, clarifying what is meant by repeat measures, etc. The authors need to consider seriously whether they wish to remain silent concerning the kind of factor analysis and rotation (if any) that they used. The authors also need to give serious attention to the issue of motivation and to whether it is automatically captured via Q methodology. Q methodology does capture perspectives, but the concept of motivation seems to imply something a bit more tricky.

Major compulsory.

**What next?:** Accept after minor essential revisions

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No