



## Commentary

### Editor:

Liam Clarke

### Submission address:

School of Nursing and Midwifery, University of Brighton, Robert Dodd Building, 49 Darley Road, Eastbourne, BN20 7UR, UK

### The jargon wheel and the total library: the problem of reliability in the research interview

In a recent commentary paper, Boyd (2008) explores issues of language and validity in qualitative research in a somewhat unusual manner. In the first part, she bemoans the decline of plain speaking and straightforward writing, declaring that she is 'increasingly frustrated by the plethora of words that obfuscates rather than elucidate', calling for 'demotic rhetoric to replace the obscure esoteric'. In her defence, she does claim that her tongue was in her cheek at the time of writing, and suggests that the relationship between language and thought might not be quite as simplistic as common sense would indicate. A whole paper, if not an entire book, could be written on this point alone, although that is perhaps for another day. In the second part of her paper, Boyd catalogues this 'obfuscating plethora' of terminology in relation to the concept of validity in qualitative research, and ends with a set of instructions for the construction of a simple 'jargon wheel' designed to produce random instances of 'validity jargon'. It is not altogether clear, however, what Boyd is attempting to achieve beyond 'a new sense of responsibility to the language' and a general warning against the obfuscating and corrupting power of 'jargon'. In the spirit of play and creativity advocated and invoked by Boyd, I have written a short response to some of the points she raises followed by a serious exploration of the topic of research reliability mediated through the playful work of the essayist, poet and short-story writer Jorge Luis Borges.

#### In defence of jargon

Interestingly, for a paper that complains about the misuse and misappropriation of words, Boyd never explores or defines the concept of 'jargon'. The dictionary offers two quite distinct meanings of the term. On the one hand it is 'the terminology or

idiom of a particular activity or group', while on the other it is 'obscure and often pretentious language' (New Penguin English Dictionary 2001). In the former case, jargon has a positive connotation and plays an important role in the communication of complex ideas between members of particular communities of practice who (to borrow from Wittgenstein) play the same language games. For example, in the very first line of her paper, Boyd uses the term 'practitioner researcher', and these words convey a particular meaning to certain readers that it would otherwise take several sentences to summarize.

Jargon is therefore more or less indispensable to all professional and academic communities, not least when it comes to keeping academic papers within word limits. However, the term 'jargon' is used mostly in the second and largely derogatory sense of obscure and pretentious language. In some cases, the obscurity and pretention is wilful and is deployed, as Boyd suggests, to confuse, colonize and/or corrupt, while in others it is employed unwittingly on the mistaken assumption that the user is being sophisticated or intellectual.

The dictionary also notes that the word 'jargon' is derived from the French for 'the twittering of birds', which again suggests two distinct meanings: to another bird, the twittering succinctly conveys a complex meaning, for example, that there is a cat around; whereas to those of us who do not belong to the community of birds, it is merely a meaningless babble and/or an aesthetic experience, depending on taste. Consider, then, the following extract from the abstract of a scientific paper:

This paper outlines a Reflexive Action Research study into the use of dreamwork in time-limited counselling, setting the work clearly in the context of Primary Care. After a brief review of the literature and a consideration of the ontological, epistemological and logical assumptions underlying the research design, the methodology is articulated in accordance with the principle of

Commentary

transparency, whereby readers can judge the trustworthiness of the account. (Boyd 2005)

Whereas a lay reader might consider Boyd's abstract to be obscure, pretentious and confusing, or perhaps as sophisticated and intellectual, or even as unintelligible twittering, to the readers of the journal *Psychodynamic Practice* it will most probably be a succinct and articulate summary of a scholarly and extremely readable paper. Furthermore, if she attempted to replace all the jargon with full explanations in plain English, the abstract would be longer than the paper it was attempting to summarize.

In the latter part of her Commentary paper, Boyd pursues the notion of jargon as meaningless twittering with the introduction of the 'jargon wheel'. The wheel, which is illustrated with a photograph, is made from cardboard and combines various words found in the literature on qualitative research validity in a variety of ways to produce random phrases such as 'pragmatic extrinsic coherence' and 'observed empirical believability' that mimic and mock some of the knots in which qualitative researchers often tie themselves.<sup>1</sup> Furthermore:

Staring blankly at the endlessness of jargon felicity and the possible trickery that underpins its constant generation . . . it became possible to see within and among the welter of words a kind of poetry, concrete in nature true, but poetic nevertheless. (Boyd 2008, p. 697)

By this point, I was beginning to believe that the whole of Boyd's paper was itself the output of some complex and sophisticated random jargon generator. My subsequent and more generous thought was that the paper might in fact be a clever attempt at reflexive parody, the point of which was, as Liam Clarke wrote in his introduction to her piece, an attempt 'to go beyond claims to postmodernist thought and to actually present something of the postmodernist motif *within the text itself*' (Clarke 2008, his emphasis). In other words, perhaps Boyd was using the example of validity jargon to make a point not just about jargon, but (albeit in an obscure way) also about the concept of validity.

<sup>1</sup>It is, perhaps, worth pointing out the interesting similarities between Boyd's jargon wheel and the 'Buzz-word Generator' proposed several years earlier by Parry & Watts (1996). In each case, three words are combined randomly to form a piece of non-sense jargon in what each text claims to be a playful way. Whereas Boyd uses a cardboard wheel, Parry and Watts have produced a table but the resulting effect is the same in both cases.

## The total library

If Boyd's 'jargon wheel' is indeed an elaborate postmodern joke designed to demonstrate the impossibility of establishing validity or truthfulness through the analysis of texts, it pales into insignificance alongside the 'Total Library' envisaged by the Argentinean writer Jorge Luis Borges. In a short story first published in 1941, Borges imagined the universe as a vast library filled with books that 'contain all the possible combinations of the twenty-odd orthographic symbols; that is, everything which can be expressed, in all languages' (Borges 1962). The vast majority of these books would contain nothing but gibberish; interminable strings of random combinations of letters such as would be produced by monkeys with typewriters. However, buried in the chaos, the library would also contain not only every book that has ever been (or ever will be) written, but also every book that is *possible* to write. Borges continues:

Everything is there: the minute history of the future, the autobiographies of the archangels, the faithful catalogue of the Library, thousands and thousands of false catalogues, a demonstration of the fallacy of these catalogues, a demonstration of the fallacy of the true catalogue . . . the veridical account of your death, a version of each book in all languages, the interpolations of every book in all books. (Borges 1962, pp. 75–76)

Clearly, Borges' Library takes us well beyond the scope of Boyd's modest 'jargon wheel', and yet each device is illustrating the same problem faced by qualitative researchers. In pushing Boyd's random validity jargon generator to its extreme and inevitable conclusion, Borges demonstrates that the problem is not, after all one of validity; it turns out that the quest for truth in qualitative research is impeded, rather, by the problem of reliability. The difficulty for qualitative researchers is not with the valid construction or discovery (depending on your epistemological and ontological stance) of truth or credibility or verisimilitude or whatever, but rather with reliably and consistently being able to recognize the truth when we see it. In the case of Borges' library, the problem is quite clear: we have no way of distinguishing between (for example) the true catalogue of the library, the thousands (if not millions) of false catalogues, the various demonstrations of the fallacy of each of these catalogues, the demonstrations of the fallacy of the fallacy of the

various true and false catalogues, and so on almost *ad infinitum*. Borges writes elsewhere that 'the fancy or the imagination or the utopia of the Total Library has certain characteristics that are easily confused with virtues' (Borges 1999). It should by now be obvious that the utopia of the Total Library holds out the promise of complete and absolute knowledge and yet delivers completely and absolutely nothing.

### The problem of reliability

For the qualitative researcher, the problem is slightly less daunting but no less extensive. In her paper, Boyd quotes from my work that 'responses are merely *different*, and that none is more or less accurate or valid than any other'. Boyd stripped the quote of its context and suggested that it might have been written 'ironically'. It was not. On the contrary, I was referring to the very real and very serious problem faced by qualitative researchers of how to make sense of diverse and sometimes contradictory information from the same interviewee. It has been known for more than half a century that the responses of interviewees will vary depending not only on their state of mind at the time, but also on who is interviewing them, the location, time of day and a multiplicity of other circumstances. It is known, for example, that a black man will respond differently to questions put to him by a white man, a white woman, a black man and a black woman (see, for example, Schuman and Converse 1971, Fugita *et al.* 2006). It has also been found that a mental health service user will respond differently to being interviewed by another service user than to a professional researcher or an academic (Clark *et al.* 1999). In some cases the variations will be slight; in others they might well produce contradictory responses.

Some writers have attempted to explain this phenomenon by suggesting that the interviewee might hold back, exaggerate or even lie in response to certain interviewers, and that the task of the researcher is therefore to minimize interviewer bias or social desirability bias (see, for example, Faulkner and Morris 2003 on the issue of mental health service users). Polit & Beck (2004) suggest that:

Interviewers are ideally neutral agents through whom questions and answers are passed. Studies have shown, however, that this ideal is difficult to achieve. Respondents and interviewers

interact as humans, and this interaction can affect responses. (p. 351)

I would suggest that the first sentence of the above statement is at best naïve and at worst simply wrong, and that the second and third sentences grossly underestimate the magnitude of the problem. Furthermore, it is founded on a positivist/realist view that the purpose of the interview is to gain unsullied access to the true and stable views of the interviewee rather than a constructionist/dialogical position whereby responses are constructed on the spot as a result of the interaction between interviewer and interviewee. From this latter perspective, the interviewer is never merely a 'neutral agent', but is an active participant in a two-way ongoing conversation comprised of speech, body language, paralanguage and a host of other social, psychological and biological cues and counter-cues that are constantly being exchanged simultaneously *in both directions*. The data generated from the interview will therefore be influenced by, among many other factors, the gender, sexuality, dress, social class, race, ethnicity, politics, and so on, of both parties.

Seen from this perspective, my so-called 'ironic' quote that 'responses are merely *different*, and that none is more or less accurate or valid than any other' is a serious assertion that there is no neutral or objective position from which to conduct a research interview, and hence no neutral or objective data to be gleaned from it. The information or research data collected from the interview will be subject to a vast and uncontrollable range of variables and will be different on each occasion. Indeed, the ever-changing dynamics throughout the course of even a single interview might produce contradictory findings.

### *Towards the total interview*

As a parallel to Borges' Total Library, we might, then, imagine the 'Total Interview' in which a single research subject is interviewed by all possible researchers in all possible locations on all possible occasions. What will be produced will be a library of all possible transcripts, some identical apart from a single word, some complimentary, some contradictory, some refuting or denying what might have been said in other interviews. The question, as it was for Borges, is thus: 'Which is the true interview?' or perhaps, 'What does this respondent *really* think?' Fortunately, perhaps, this position is never encountered in the real world. Generally,

researchers only have time to conduct a single interview with each respondent. What is remarkable, however, is how they are willing to accept the findings from this random sample of one transcript from a near-infinite library of possible transcripts as revealing something meaningful, something that has a certain *validity*, something that points towards the truth, however it might be conceived. In other words, the reliability of this single random interview situation tends to be accepted almost without question, or at the very best, the researcher strives to minimize the effects of bias to the extent that they are thought to pose no real threat to the data. There seems to be no desire to face up to the fundamental and built-in flaw at the heart of the interview method of data collection; that the same question posed to the same interviewee under different circumstances by different researchers on different occasions might elicit very different and perhaps inconsistent or contradictory findings.

There is, as far as I can make out, no solution to this problem of reliability. The interview can, of course, be repeated under a variety of different conditions in the hope that similar data emerge and that the study therefore has 'test-retest reliability', or a colleague could be asked to conduct an interview with the respondent to demonstrate 'inter-rater reliability', and significant similarities might be found in the transcripts. As Borges pointed out in relation to the Total Library:

Each book is unique, irreplaceable, but (inasmuch as the Library in total) there are always several hundreds of thousands of imperfect facsimiles – of works that differ only by one letter or one comma. (Borges 1962)

In the Total Interview, as in the Total Library, the discovery of a second, similar text implies nothing about the validity, accuracy or truthfulness of the first, any more than the discovery of a contradictory text can be assumed to negate it. In a world where every research encounter produces unique and singular information about how to practice mental health nursing, there is at least some comfort in the thought that, in any case, every therapeutic encoun-

ter between every patient and every nurse is also unique and singular. Perhaps, then, the only valid conclusion to be drawn is that far too much emphasis is placed on reliable research evidence as being able to assist practice in an inherently unreliable world.

G. ROLFE RMN PhD

*Professor of Nursing*

*School of Human and Health Sciences*

*Swansea University*

*Singleton Park, Swansea, UK*

## References

- Borges J.L. (1962) The library of Babel. In: *Fictions* (ed Kerrigan, A.), pp. 72–80. Jupiter, London.
- Borges J.L. (1999) *The Total Library: Non-Fiction 1922–1986*. Penguin, Harmondsworth.
- Boyd J. (2005) Where two worlds meet: an exploration of the client's experience of dream-work in time limited therapy. *Psychodynamic Practice* 11, 189–204.
- Boyd J. (2008) Commentary. *Journal of Psychiatric and Mental Health Nursing* 15, 695–698.
- Clark C., Scott E., Boydell K., *et al.* (1999) Effects of client interviews on client reported satisfaction with mental health services. *Psychiatric Services* 50, 961–963.
- Clarke L. (2008) Commentary. *Journal of Psychiatric and Mental Health Nursing* 15, 694.
- Faulkner A. & Morris B. (2003) Expert paper: user involvement in forensic mental health research and development. *NHS Programme on Forensic Mental Health Research and Development*, University of Liverpool.
- Fugita S.S., Wexley K.N. & Hillery J.M. (2006) Black-white differences in nonverbal behavior in an interview setting. *Journal of Applied Social Psychology* 4, 343–350.
- New Penguin English Dictionary*. (2001) Penguin Books, Harmondsworth.
- Parry G. & Watts F.N., eds (1996) *Behavioural and Mental Health Research*. Taylor & Francis, Hove.
- Polit D.F. & Beck C.T. (2004) *Nursing Research: Principles and Methods*, 7th edn. Lippincott, Philadelphia, PA.
- Schuman H. & Converse J.M. (1971) The effects of black and white interviewers on black responses in 1968. *Public Opinion Quarterly* 35, 44–68.