Questionnaires in Dictionary Use Research: A Reexamination

Robert Lew
Department of English, Adam Mickiewicz University
Niepodleglosci 4
PL-61874 Poznan
Poland
rlew@amu.edu.pl

Abstract
The present paper re-examines the usefulness of questionnaires in dictionary use research, using Glynn Hatherall's well-known criticism of questionnaires as a starting point. It is argued here that charges directed at questionnaires apply equally easily to the alternatives suggested by Hatherall. It is claimed that some research questions require a questionnaire approach. It is also demonstrated through example that unreliability of questionnaire-based studies may well result from design factors unrelated to questionnaires themselves. Use of multiple methods and careful design is advocated. Finally, suggestions are offered for improving questionnaire design in dictionary use research.

Criticism of questionnaires

Glynn Hatherall denies the usefulness of questionnaires in dictionary use research in the following words: "Are subjects saying here what they do, or what they think they do, or what they think they ought to do, or indeed a mixture of all three?" [1984:184]. This often-quoted passage did much to make questionnaires an unpopular – if not downright deprecated – instrument among metalexicographers.

As an alternative to questionnaires, Hatherall advocates direct observation as "the only reliable method of collecting data on dictionary user behaviour" [1984:184]. Broadly speaking, the thrust of Hatherall's criticism of questionnaires is that they do not measure what they purport to measure: it concerns the *reliability* of measurement. Yet this is also a charge that can be (and has been) justifiably made against direct observation techniques (observer's paradox). Hatherall himself appears to be aware of this problem: "Under such conditions [while being observed - RL] it would probably be difficult for the subjects to behave normally" [Hatherall 1984:184]. Hatherall also rightly points out that direct observation is a rather inefficient technique of data collection.

The technique proposed by Hatherall is based on written self-report protocols. In Hatherall's [1984] study, subjects were asked to record on a special form details of all dictionary searches they performed while translating, and their level of satisfaction with the information found. In view of Hatherall's methodological concerns, it is somewhat surprising that Hatherall does not address the issue of measurement reliability of this technique. Subjects are simply expected to behave "in a normal way" [Hatherall 1984:185] while working on their difficult translation tasks (here native speakers of English translating a difficult business text from English into German) and, at the same time, monitor all their activities and complete report forms. It appears to be a risky assumption indeed that the task of translating a difficult text into a foreign language remains unaffected if a portion of the

subjects' attention and cognitive potential is reassigned to the novel and complex task of monitoring and recording every step of their activities. Quite the reverse, it seems reasonable to assume that under such conditions subjects would behave anything but "in a normal way".

The right method?

The preceding discussion may suggest that, as no single method is without its problems, restricting dictionary use research to just one specific method or technique is not the best policy to follow. As *all* methods have their limitations, it is part of researchers' job to carefully identify these limitations and try to compensate for them as much as possible. One of the best ways of achieving this is to use a variety of methods and compare their results (through what is sometimes known as *methodological triangulation*, cf. Nesi [2000:12]; cf. also Hartmann's [2001:115] *methodological plurality*). This tendency is rather clearly reflected in the recently published book-length contributions dealing with dictionary use such as Atkins [1997] or Tono [2001] as well as in other recent publications in the field (for which see Dolezal and McCreary [1999]).
It might also be pointed out that there exist aspects of dictionary use which appear to present serious challenges to direct observational techniques. Here one might just mention issues related to long-term and historical aspects of dictionary-related behaviour (e.g. age when first dictionary was purchased, frequency of dictionary use at home, etc.), and attitudinal aspects of dictionary use (e.g. dictionary preference, satisfaction with dictionaries, etc.).
There is, of course, no space here to consider every imaginable research method that a dictionary use researcher might choose to use. In what follows, I will concentrate on questionnaires.

The questionnaire is not everything

Hatherall [1984] founds his criticism of questionnaires on a comparison of results from two questionnaire studies [Béjoint 1981; Hartmann 1982], pointing to the differences between the two studies with regard to subject-reported frequencies of the types of words most often looked up in dictionaries. Hatherall appears to suggest that the differences between the results of the two studies should be interpreted as indicative of unreliability of questionnaire-based techniques. In doing so, Hatherall disregards the fact that the two studies are based on different groups of subjects (French university students vs. British learners) using different types of dictionaries (monolingual vs. bilingual) for different target languages (English vs. German). Given these substantial differences in sampling and conditions, it would be in fact surprising if convergent profiles of dictionary use had been obtained in the two studies – by *any* method. One would naturally expect to get different results from the two studies for reasons related to any and all of the following:

- different national lexicographic traditions
- difference in foreign language competence level
- difference in age
- different types of dictionaries
- different target languages

Furthermore, the measures reported in the two studies are – as far as I can tell – not equivalent and thus not directly comparable. If I understand the original reports correctly, then Béjoint [1981] reports only the frequency of the modal value, i.e. the most common

response for a given type of word, ignoring the influence of all the other responses, whereas Hartmann [1982] gives a mean percentage measure (counting all responses).
Thus, for all we know, there is really no evidence that the reported differences are indeed (as suggested by Hatherall) an artefact of using questionnaire methods, and that they are not in fact valid findings.

Replications and comparisons

The two studies by Béjoint [1981] and Hartmann [1982] were not originally designed to be directly comparable. The two authors may have had different research questions, and this may have been reflected in the format of the questionnaires and the phrasing of the questions. But, again, the fact that results of two studies may not be comparable even if the same broad method is used is hardly unique to questionnaire-based studies. Exactly the same will hold for experimental or direct observation studies.
If the aim is to produce a new study that is comparable with an existing one, than the methodologically correct option would be to exactly replicate – as far as possible – the instruments and conditions of the original study, except any conditions whose influence might be of direct interest.
A good example from questionnaire-based research on dictionary use is offered by two studies by Quirk [1974] and Greenbaum, Meyer & Taylor [1984], respectively. The latter study was specifically designed as a replication of Quirk's study, and its primary aim was to compare responses by US students with those of their British counterparts obtained by Quirk [1974]. It is then methodologically commendable that Greenbaum, Meyer & Taylor set out to "give American College students a set of questions similar to those that Quirk asked the British students" [1984:31] (with the understanding that the degree of similarity was substantial). Consequently, any differences between the responses reported in the two studies are interpreted by Greenbaum, Meyer & Taylor [1984] as reflecting underlying British - American differences.
However, a careful reading of the two articles reveals that they differ quite substantially in the structure of their samples: the British subjects were all "half way through the first year of their studies" [Quirk 1974:150], while the majority of American subjects were in their third or fourth year [Greenbaum, Meyer & Taylor 1984:31]. Greenbaum, Meyer & Taylor's unfortunate choice of a sample so radically different from that of Quirk's original study was probably motivated by convenience, and that may be understandable (though a smaller sample of freshmen alone may still have been preferable). What is more difficult to understand is why Greenbaum, Meyer & Taylor completely ignore the difference in sample structure in their discussion of results. At least some of the effects interpreted by Greenbaum, Meyer & Taylor as British - American differences appear to be quite likely to be sensitive to year of study (e.g. difficulty in understanding definitions; comprehensibility of phonetic transcription), but no such interpretation is even considered. With the American subjects coming from all four years of the undergraduate programme, a more balanced comparison would be obtained if year of study was entered as a predictor in the analysis.
The above examples demonstrate that the reliability of questionnaire-based research on dictionary use does not depend solely on the questionnaires used: due attention should be paid to *all* aspects of the design. However, a well-designed questionnaire is certainly a key element, and I would like to conclude this paper with a few practical suggestions in this area.

How to make better questionnaires

In my view, researchers using questionnaires (or, better still, *planning* to use questionnaires) would be well advised to consult existing questionnaire design manuals written by experts in sociometry and psychometry [Berdie & Anderson 1974; Bradburn, Sudman & Blair 1979; Oppenheim 1992; Sudman & Bradburn 1982]. Even though the aims of metalexicographers are qualitatively different from those of most sociologists and psychologists, surely some established standards of good practice in the design and application of questionnaires can be found that are relevant and potentially beneficial to questionnaire-based lexicographic research.

That said, I would still like to offer here a short list of *dos* and *don'ts* that I have arrived at on the basis of my experience with questionnaire design for dictionary use research and from critical scrutiny of published and unpublished questionnaire-based studies, in the hope that the list can serve as a quick checklist, especially, but perhaps not exclusively, with newcomers to the field.

Do:

- write your questionnaire in the subjects' native language
- pay attention to clean, unambiguous graphical layout
- consider in each case whether multiple choice or open-ended or mixed question format is most appropriate
- decide before the design is complete how the results will be coded and processed
- screen your questions and multiple-choice answers for possible bias
- ask a colleague or two to read through a draft of your questionnaire
- pilot your questionnaire
- allow appropriate time for your questionnaire to be completed

Don't:

- use technical language that subjects might not understand
- use complex syntax
- use negatives in questions
- let page breaks split questions
- put nonessential questions in the questionnaire just because others had them
- give away your own position or preference in any way

Conclusion

There is certainly place for questionnaires in dictionary use research. Other methods should not supplant but rather supplement questionnaires. Ultimately, questionnaires are only as good – or bad – as you make them.

References

[Atkins & Varantola 1997] Atkins, B.T.S. & K. Varantola. 1997. Monitoring dictionary use. *International Journal of Lexicography* 10 (1): 1-45.

[Béjoint 1981] Béjoint, H. 1981. The foreign student's use of monolingual English dictionaries: A study of language needs and reference skills. *Applied Linguistics* 2 (3): 207-222.

[Berdie & Anderson 1974] Berdie, D.R. & J.F. Anderson. 1974. *Questionnaires: design and use.* Metuchen, N.J.: Scarecrow Press.

[Bradburn, Sudman & Blair 1979] Bradburn, N.M., S. Sudman & E. Blair. 1979. *Improving interview method and questionnaire design. The Jossey-Bass social and behavioral science series.* San Francisco: Jossey-Bass.

[Dolezal & McCreary 1999] Dolezal, F.T. & D.R. McCreary. 1999. *Pedagogical lexicography today. A critical bibliography on learners' dictionaries with special emphasis on language learners and dictionary users. Lexicographica Series Maior.* Tübingen: Niemeyer.

[Greenbaum, Meyer & Taylor 1984] Greenbaum, S., C.F. Meyer & J. Taylor. 1984. The image of the dictionary for American college students. *Dictionaries: The Journal of the Dictionary Society of North America* 6: 31-52.

[Hartmann 1982] Hartmann, R.R.K. 1982. Das zweisprachige Wörterbuch im Fremdsprachenerwerb. *Germanistische Linguistik* 3-6 (80): 73-86.

[Hartmann 2001] Hartmann, R.R.K. 2001. *Teaching and researching lexicography.* Harlow: Pearson Education Limited.

[Hatherall 1984] Hatherall, G. 1984. Studying dictionary use: Some findings and proposals. In: Hartmann, R.R. (ed.), *LEXeter '83 Proceedings: Papers from International Conference on Lexicography at Exeter, 9- 12 Sept. 1983, Lexicographica Series Maior 1,* 183-189. Tübingen: Niemeyer.

[Nesi 2000] Nesi, H. 2000. *The use and abuse of EFL dictionaries.* Vol. 98. *Lexicographica Series Maior.* Tübingen: Niemeyer.

[Oppenheim 1992] Oppenheim, A.N. 1992. *Questionnaire design, interviewing, and attitude measurement.* London - New York: Pinter Publishers.

[Quirk 1974] Quirk, R. 1974. The image of the dictionary. In: Quirk, R. (ed.), *The linguist and the English language,* 148-163. London: Edward Arnold.

[Sudman & Bradburn 1982] Sudman, S. & N.M. Bradburn. 1982. *Asking questions: a practical guide to questionnaire design. Jossey-Bass series in social and behavioral sciences.* San Francisco: Jossey-Bass.

[Tono 2001] Tono, Y. 2001. *Research on dictionary use in the context of foreign language learning : focus on reading comprehension. Lexicographica Series Maior 106.* Tübingen: Niemeyer.