

QUALITY IN INTERPRETING: SOME METHODOLOGICAL ISSUES

By
Barbara Moser-Mercer
ETI, Université de Genève

1. Introduction

Quality in interpreting is a frequently discussed yet little explored concept. As services are increasingly evaluated the way products are, with ISO 9000 and TQM having become the hallmarks of production in post-industrial societies, translation and interpreting cannot claim special status as “works of art”, to be evaluated much the same way as gymnastics, dance or figure skating competitions. This is not to say that translating or interpreting lack artistic value or creativity, but current practice puts them squarely into the category of services; as such they become subject to performance evaluations similar to those carried out on products and other types of services. Especially interpreting, which in the eyes of the consumer carries a hefty price tag, is becoming the object of such evaluations. There is no point in lamenting the apparent loss of special status. As professional skills have evolved since antiquity, and particularly since the Nuremberg trials - seen by most as a historical turning point with the introduction of simultaneous interpreting - the public is coming to expect more and more of interpreters, while at the same time judging their performance often too harshly. This is frequently done without bearing in mind the multitude of factors that can affect the quality of interpreting.

2. Defining quality in interpreting

Before we can speak of evaluating quality we first need to define the product or service and then the notion of quality. The definition chosen for this article will be for conference interpreting, as this corresponds to the author’s main area of experience. It is interesting to note that from a professional perspective the notion of interpreting has not been defined, whereas there exists a commonly approved definition for the interpreter.

A conference interpreter is a qualified specialist in bilingual or multilingual communication. He/she makes this communication possible between delegates of different linguistic communities at

conferences, meetings, negotiations or visits, where more than one working language is used, by comprehending the concepts of the speaker's message and conveying them orally in another language, either in consecutive, simultaneous or whispering.

Besides carrying out a thorough preparation of the subject and terminology, a conference interpreter must possess a wide general knowledge in order to deal with all matters under discussion.

Conference interpreters are, moreover, bound to respect the code of professional ethics, including the strictest professional secrecy. (AICC, 1994)

Based on this definition we can now proceed to develop the notion of quality. Since quality tends to be a relative notion, with beauty often in the eyes of the beholder, it appears difficult to agree on an absolute definition of what should be considered quality in interpreting. The large variety of working conditions that all potentially influence quality makes it necessary to introduce certain caveats.

I propose to work with the notion of **optimum quality**:

Optimum quality in professional interpreting implies that an interpreter provides a complete and accurate rendition of the original that does not distort the original message and tries to capture any and all extralinguistic information that the speaker might have provided subject to the constraints imposed by certain external conditions.

The caveat **external conditions** requires further elaboration as it obviously describes the constraints interpreters often have to work under and which may interfere with their providing outstanding quality in an abstract sense. Hence the following addition to the above definition seems appropriate:

Optimum quality is the quality an interpreter can provide if external conditions are appropriate.

These conditions include the following:

- physical environment (booth dimensions, equipment, air quality, position of booths, lighting, etc.),
- complexity of subject matter under discussion,
- change in subject matter,
- adversarial nature of meetings,
- discourse characteristics (density of text, emotionality, coherence, etc.),
- delivery (speaking speed, accents, graphics, presentation, etc.),

- preparation of interpreter (documentation),
- team size, length of turn, load during working day, number of consecutive meetings,
- speakers speaking simultaneously,
- interpreter's emotional response,
- competence and availability of technician,
- etc.

This list can obviously be continued, as most interpreters can provide additional variables from their own experience. The idea is, however, not to provide an exhaustive listing, but to emphasize that when evaluating quality of interpreting services no abstract approach will be appropriate, but quality will always have to be evaluated against the background of the working conditions that prevail in the particular situation under observation.

Next it would seem important to return to the question of whether interpreting is an art, a work of art, a product or a service. If interpreting were an art, we would probably want to evaluate it like a figure skating competition. If it were a work of art we would submit it to art critics, if it were a product we could focus on the output and would have to disregard a number of those factors listed under external conditions above. It would seem to me that interpreting should be categorized as a service that is provided to those who do not understand a message delivered in the original language. This broadens the notion of interpreting and ensures that quality is not seen in the narrow and abstract sense as something only inherent in the message delivered.

3. Finding out about quality in interpreting

When speaking of quality we immediately think of how we can measure it, or evaluate it, or assess it. Which term to use will depend on the very reason why we decide to take a look at interpreting quality: if we are concerned with the quality of interpreting in a natural setting, we probably want to evaluate it. If, however, our goal is to analyze the product of interpreting in a laboratory setting, we would most likely proceed to measurement. In working with student interpreters, novices on their way to becoming experts, neither evaluation nor measurement will do, but assessment would be the proper approach, until, at least, the student's final course exams, when we are more likely going to combine both evaluation and assessment.

The reader may feel that such fine distinctions serve more to confuse than to clarify the issue at hand, that every professional interpreter knows intuitively what quality is and could clearly distinguish between good and bad quality. I would agree that most interpreters have some intuitive feel for quality, but argue

that this fails him or her more often than he or she likes. Towards the end of a turn interpreters do not seem to be able to evaluate accurately the quality of their interpretation, and awareness of poor quality develops only long after a significant decline has occurred (Moser-Mercer et al., 1996).

A thorough examination of the notion of quality in interpreting requires looking at it from a variety of different perspectives:

- Quality cannot be seen only through the eyes of the interpreter, the provider of the service who bases himself on his own set of self-imposed standards.
- The quality of the service performed can also be evaluated in terms of the judgments, needs and expectations of users of that service.
- Intermediaries (employers, agencies, etc.) may introduce yet additional variables that might change the way quality of one and the same interpretation may be evaluated.
- There are also interpreting colleagues whose evaluations may play a crucial role in the career development of a budding interpreter.
- The experimenter may be interested in finding out how quality changes when an independent variable is manipulated.

Next it would seem to be important to define the purpose of quality evaluation: Does the interpreter want to know how many mistakes he made during any particular turn, or does he want to know whether his performance is better during technical meetings than when more philosophical issues are discussed, or whether his new C-language is good enough to really be called a C-language; or is he interested in finding out whether he works better from B to A than from A to B during any particular meeting. Similarly, intermediaries may have different sets of questions to ask with respect to their notion of quality: Does this new interpreter they recruited really sound as smooth as the other, well-known members of the team; have the interpreters really studied their terminology; what happens if the interpreters do not manage to get the documents - will they really flounder? The booth mate might be interested in just how his colleague dealt with a given idiom or play on words; how his colleague managed in the face of an almost impossible situation with a speech being read with a heavy regional accent and at top speed with no document for the interpreter to rely upon. The experimenter may wish to manipulate input speed, length of turn, type of material presented in order to observe quality as the dependent variable.

Needless to say, few of the above scenarios lend themselves to exactly the same research design, yet they all have to do with analyzing quality. Obviously, it cannot be in our interest to advocate one way of analyzing quality, given the multitude of perspectives, settings and questions researchers are interested in.

Our next task should then be to provide some idea as regards possible research designs for some of the questions that researchers might be interested in. It is impossible to exhaust all the possibilities in a short article, but I would like to develop some guidelines for choosing an appropriate method.

We need to return briefly to our distinction between quality evaluation, quality measurement and quality assessment. We said that evaluation was something a researcher would engage in when trying to obtain information about the quality of a service. Such an approach would provide answers to such questions as: Did the interpreter grasp the essential thrust of the speech? Was the user satisfied with the delivery? Did he feel the interpreter used proper technical terminology? Is the interpreter aware of the mistakes he made? And we could go on to list a large number of questions that a researcher might ask in order to **evaluate the quality of interpreting in the field**.

Questions such as: How many mistakes did the interpreter make during this particular turn, or while interpreting a particular speech? To what extent does the density of the original affect the quality of the interpretation? What is on average the number of figures an interpreter gets wrong while interpreting a treasurer's report? What effect does preparation have on the quality of the interpretation? - are best answered by setting up **laboratory experiments** during which **quality is measured** using specific types of scales.

If we wish to find out how a particular pedagogical tool affects qualitative improvement in students' performances, or how much practice is needed for a student to progress from one specified level to the next, we will **assess** changes in **quality** over a given period of time, usually in a **classroom**, a **semi-controlled** or a **laboratory setting**.

Some of the problems we can discern in studies on quality in interpreting are a direct consequence of the failure to make precisely these distinctions, where methods are chosen that are not appropriate to the questions asked. It would also seem to me that some studies on quality fail to state clearly what questions the researcher wants to have answered, simply because the notion of quality is not clearly defined and the researcher just wants to find out "about quality in interpreting", a term that as we have seen above requires further definition. The design of the study (whether survey, experimental or other) is thus chosen without reference to any clear criterion (evaluation is used instead of measurement, for example), and the study then has little or no face or content validity. In other words, one cannot deduce from the data obtained any reliable information about the question asked.

We have tried to clarify first of all the perspective. Next we need to ensure that we chose the right method that will help us answer the questions we wish to put.

3.1 The user perspective

Based on what we have learned from several user surveys (Kurz, 1989; 1993; Bühler, 1986; Kopczinski, 1994; Moser, 1995) users of interpreting services appear to focus mostly on the following variables: accuracy of content, terminological accuracy, simultaneity of delivery, rhetorical skills of the interpreter, voice and microphone discipline.

It appears obvious that one of the best ways of finding out whether a user is satisfied with a particular service is to ask questions. The methods of empirical social science research in the form of interviews and surveys thus lend themselves best to this approach.

Questioning methods can be *formally structured* with every respondent receiving exactly the same questions, most likely in the same order, or they can tend towards the *unstructured*, which allows the researcher to be flexible towards the respondent with a good chance of obtaining genuine, more complete answers.

The *structured* interview will be quick to administer, is easily replicated, yields generalizable results, allows for relatively simple data analysis, with rather low influence from interpersonal variables; it is also highly reliable. On the other hand, the structured interview has its drawbacks in that the respondents are completely constrained by the questions, that the information gained is rather narrow and may be distorted by ambiguous, complex wording and an inappropriate response choice list, in short it suffers from all the difficulties associated with questionnaires.

The *unstructured* interview can be moulded to the individual situation and context and yields richer, fuller information. Interviewees usually feel more relaxed and not under pressure of assessment. The disadvantages are that an *unstructured* interview is unsystematic and thus different information is gathered from different people. It can thus be difficult to analyze the variety of information gathered.

One way to exploit the advantages of the *unstructured* interview is to keep the procedure informal, not to ask pre-set questions in quite the same way every time and to provide interviewers with instructions that contain an outline of the topics and issues to be covered and questions to be asked. The interviewer is free to decide how to present questions and how to phrase them. This increases the consistency of information and facilitates data analysis; the information obtained remains genuine while the interviewer maintains flexibility. The information is rich and realistic. Of course, different wordings may shift the emphasis and allow for different interpretations of a question. The interviewer may also miss important topics. Thus the results are perhaps less generalizable. One way to counteract these disadvantages is to include more open-ended questions.

In all of the above types of interviews the interviewer should avoid asking complex, ambiguous, leading or emotive questions, to ask two or more questions at once, to ask “why?” questions that waste everyone’s time or questions that require simply a yes or no answer, thus yielding only minimal information.

Data can be recorded either by taking notes and/or by recording on audio or on videotape.

In order to guarantee smooth functioning of the interviews, pilot interviews should be conducted and evaluated. These will allow researchers to catch any flaws in interview technique and study design.

When conducting *surveys* researchers will ask a lot of people a lot of questions in order to obtain an accurate description of what users think. Data collected that way can serve a purely descriptive purpose. This often represents the initial stage of research in finding out about quality. On the basis of data collected one can formulate certain hypotheses about the way quality is perceived by the user and test these by means of an additional survey or with in-depth interviews. When embarking on larger-scale surveys it is important to decide carefully on the type of *sample* (if we wish to find out about how users perceive quality it is no good to survey interpreters or students), the *mode of questioning* (face-to-face or in written form), and the *questions* themselves. If the written mode is used the questionnaire must be exceptionally clear and the instructions unambiguous. One also needs to bear in mind that the return rate is usually rather low; by contrast, the rate at which an oral interview would be refused in a clearly defined sample is much lower. The low return rate also raises the question as to whether those who did not return the questionnaire are neutral towards the issue or just could not be bothered; with return rates of between 10 and 25%, that difference in opinion does matter when the researcher proceeds to generalize from the data obtained. (Pöchacker, 1994: 330)

As to the questions asked there are some general rules. First, one should ask only for the minimum of information required for the purpose of the study and avoid those that might seem interesting but have little or no bearing on the purpose of the study. Respondents must be able to answer a question. When interested in comparing the quality of simultaneous and consecutive interpreting one needs to ensure that the respondent has had occasion to use both. One should also abstain from jargon; terms such as A-language, or even consecutive may mean little to the uninitiated. Questions ought to be phrased in such a way that they can be answered truthfully. Respondents may prefer to base their answers on widely held opinions, or on what they believe the research director would like to read, than on what s/he really thinks.

When answers to a question are to be given on a rating scale the questionnaire designer needs to ensure that the same attitude scales are used

throughout, and that the scale has an equal number of positive and negative statements. This will ensure that during data analysis correlations can be calculated across questions. It is also advisable to pilot scales, either as part of piloting the entire questionnaire, or as individual questions. Even an unambitious questionnaire constructed by students should be piloted at least once to avoid problems at the data analysis and interpretation stage.

A word about the number of questions included in a questionnaire and their organization: Obviously, we don't wish to ask too many questions, but there should be enough to achieve reliability and to ensure that random errors cancel each other out. When organizing the questions one should avoid producing a potentially long, uninterrupted series of yes or no answers as this will influence the respondent in a certain direction and could distort answers to later questions.

Whether one chooses the interview method or the questionnaire method, or a mixed design, one still needs to ensure that the method is valid, i.e. that it measures what the researcher intended to measure. The simplest way is to look at the contents of a questionnaire or at the interview questions, which is best done by someone other than the investigator himself. The investigator also needs to check for reliability of the questionnaire or interview questions used. Several pilots could help achieve certainty in that regard, the other method being a repeat investigation on a different occasion. Lastly, the issue of which population is chosen for the study deserves attention: students of interpreting will most likely have an approach to quality that is different from that of a user, or from that of an employer. Thus, the results of studies designed with students as the target group cannot be generalized to the population of users at large. So if we want to find out about quality in interpreting as seen by the user of this service it serves no purpose to sample students, or interpreters for that matter, unless the research question is: Do students of interpreting (or interpreters) hold different opinions with respect to quality from actual users?

3.2 Employer's perspective

Virtually no systematic research has been carried out with respect to what employers expect from interpreters in terms of quality. Thus the following variables are merely suggestions as to how one could proceed to investigate the matter. Prospecting and pilot interviews will certainly yield the necessary information that can then be used to guide the final research design.

It would seem that accuracy of content and technical terminology, rhetorical skills, but also team discipline, adaptability to different situations, flexibility in scheduling, price, availability and loyalty to a particular employer may all enter into the employer's perception of quality.

From a methodological point of view the methods described for investigating users' perceptions of quality can all be used to investigate employers' opinions. It is important to bear in mind parameters such as different employer categories (agencies, PCOs, large language services of international organizations), different types of conferences, and different cultural settings.

3.3 Measuring quality for experimental purposes

While the purpose of assessing users' and employers' perceptions of quality is to obtain a better idea of what these groups are looking for in an interpreter's performance and whether the service can be considered satisfactory, laboratory experiments may be conducted for a variety of purposes, with quality being the dependent variable.

Quality may be measured in order to find out how long an interpreter is capable of working before performance drops, how different settings (different types of conferences) impact on an interpreter's performance, how the interpreter's service changes over his life span, what differences there are between interpreting into an A and interpreting into a B language, etc. In all of these research scenarios quality is the dependent variable to be measured usually under laboratory conditions. Here, the original (as available from a transcript) is to be compared to the interpretation (which must also be available as transcript).

The experimenter must carefully control the input (level of difficulty of text through propositional analysis, rating of text, or similar reliable means; speed at which it is delivered; subject matter; etc.), the interpreters (years of experience, language combination, language direction, etc.) and the setting (technical infrastructure).

Measuring quality without *error scales* is impossible. Such scales have been developed by several researchers over the years and have prompted a great deal of debate (Barik, 1972; 1973; Gerver, 1969). If we return to the earlier definition of optimum quality it would appear that accuracy and completeness rank high, and an error scale should thus not only list different types of errors, but also rank and weigh them appropriately. Each error type should be clearly defined and there should be no overlap between different error types. Furthermore, researchers need to guard against too fine a distinction among error types, as this will make the scale cumbersome to use and most likely yield too few data per category to allow for meaningful interpretation of results.

Anyone who is familiar with correcting student translations knows that no two versions are alike and no two teachers will correct exactly the same way. It thus should be no more than common sense for any investigator to plan on having several judges rate the quality of an interpretation and to proceed to an

analysis of interjudge reliability before including the results of error analysis in further data analysis. Error scales need to be piloted with different interpreting samples and different judges. Based on my own experimental experience I recommend that judges be interpreters who are familiar with interpreting strategies and will thus, for example, not count every omission as omission without having reassured themselves that the interpreter has not included the "missing" information somewhere else, or provided an elegant synthesis. Sometimes it might be of interest to play back the interpretation to the interpreter to obtain additional information on strategies used. Such introspective data can be used to complement the more objective approach of error analysis.

Error analysis is time-consuming. It is the dirty work that must be done before one can proceed to more interesting analyses and hypothesis testing. It thus pays to invest in it and to employ the best possible scales. If the study looks at interpreters' strategies and the interpreting process additional tools are needed, such as software that will provide a visible representation of the original and the interpretation. Since we are dealing with overlapping speech, a specialized software tool is needed that will mark up and annotate the corpus. For a review of such tools see Meyer et al. (1994), Rehbein et al. (1994), and Edwards & Lampert (1993). These packages run on ordinary PCs and are, as far as price is concerned, certainly within the reach of any university department and even that of the individual researcher.

3.4 Quality and ongoing student assessment

We don't usually think of quality when assessing students, it is nevertheless what we are implicitly looking for. As novices move on to become experts we are constantly assessing them along the lines of the only parameter that seems meaningful: the quality of the interpretation as an indicator of a student's progress.

While teachers don't usually think of bringing sophisticated methods of assessment into the classroom (aptitude tests and final exams are the exception), as judges of quality they need to be aware of a number of variables.

The quality of student performances cannot be assessed in a vacuum, (something we already know from our definition of optimum quality). Just like researchers, teachers need to know what they are looking for, what level of quality they are expecting. Thus, they need to bear in mind overall curriculum design, syllabus design, the level of expertise of the trainee and the goal of the class during which assessment is carried out. It seems obvious that a novice will render quality different from that of an intermediate or advanced student. Teachers must be aware of what a sensible progression within the constraints of the

overall curriculum looks like and, as a consequence, set realistic goals for the trainee. Often teachers expect and encourage attention to detail over meaning in the beginning of a training program, and as a result, obtain detail without meaning, since it is well known from studies in expertise (Gordon, 1992; Künzli & Moser-Mercer, 1995) that novices move from micro to macro context and from a molar to a global understanding of discourse. Teachers should also be familiar with the interpreting process, i.e. having reflected upon how one proceeds, how one develops strategies and automatisms and how one acquires the world knowledge and technical knowledge necessary for expert performance.

Since interpreting classes usually have a restricted number of students, teachers should be able to develop individual student profiles. In order to get started one can ask students to keep a diary on their performance in interpreting classes over the space of a few weeks. These are then analyzed by the teacher with relevant information as to students' self-perceived strengths and weaknesses together with relevant information from the teacher's diary for the same student for the same period entered into a matrix. Teachers can then enter regular assessments into the cells of the matrix, with parameters on the vertical line and time charted on the horizontal line. As time goes by the cells will be filled with observations and will give both teacher and student a good overview of where the trainee stands at any particular point in time. Once teachers have collected such information for several years they develop a much better feel for whether or not a trainee is on target. As indicated above, assessment must correspond to the level of progression stipulated in the curriculum and in the syllabus. This avoids overcorrection, which can be demotivating, and undercorrection, which can raise false expectations.

As we have seen, there is no need to resort to error scales or surveys or questionnaires when it comes to assessing quality of student performances, but a systematic approach is still needed to ensure that the teacher assesses what he should.

4. Conclusion

While this article has tried to explore how quality can be looked at from different vantage points and how it can ultimately be evaluated, measured or assessed, it by no means claims to be an exhaustive treatment of the subject. It is intended as a guide for those who are interested in researching quality in interpreting, whether they are students, professional researchers or professional interpreters, and who are looking for a definition of the notion of quality and for methodological advice on how to research it. Sound research methodology should never be reserved to professional research, but must be applied to even

the simplest study carried out by students. Every study must be clear on what question it wishes to answer, which population is targeted, and what method is chosen for data collection and analysis. Every study must demonstrate validity and reliability, or put another way:

Scientific research needs to be seen for what it truly is; a way of preventing me from deceiving myself in regard to my creatively formed subjective hunches which have developed out of the relationship between me and my material. (Rogers,1961)

Bibliography

- Abelson, R.P. (1995): *Statistics as Principled Argument*, Hillsdale, NJ.: Lawrence Erlbaum Associates.
- AIIC (1995): *Advice to Students Wishing to Become Conference Interpreters*, Geneva, AIIC.
- Bühler, H. (1986): "Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters", *Multilingua*, 5(4), pp. 231-235.
- Coolican, H. (1994): *Research Methods and Statistics in Psychology*, London, Hodder & Stoughton.
- Edwards, J.A. & Lampert, M.D. (1993): *Talking Data: Transcription and Coding in Discourse Research*, Hillsdale, NJ, Lawrence Erlbaum Associates.
- Frey, G. (1970): *Philosophie und Wissenschaft*, Stuttgart, Kohlhammer.
- Gadenne, V. (1976): *Die Gültigkeit psychologischer Untersuchungen*, Stuttgart, Kohlhammer.
- Gile, D. (1995a): *Regards sur la recherche en interprétation de conférence*, Lille, Presses Universitaires de Lille.
- Gile, D. (1995b): "Fidelity assessment in consecutive interpretation: An experiment", *Target*, 7(1), pp. 151-164.
- Gordon, S.E. (1992): "Implications of cognitive theory for knowledge acquisition", in R.R. Hoffman (ed.), pp. 99-120.
- Greene, J. & D'Oliveira, M. (1994): *Learning to Use Statistical Tests in Psychology*, Philadelphia, Milton Keynes.
- Hoffman, R.R. (ed.)(1992): *The Psychology of Expertise: Cognitive research and empirical AI*, New York, Springer.
- Hoffman, R.R.; Shadbolt, N.R. & Burton, A.M. (1995): "Eliciting knowledge from experts: A methodological analysis", *Organizational Behavior and Human Decision Processes*, 62 (2), pp. 129-158.

- Kopczynski, A. (1994): "Quality in conference interpreting: Some pragmatic problems", in S. Lambert & B. Moser-Mercer (eds.), *Bridging the Gap: Empirical Research in Simultaneous Interpretation*, Amsterdam, Benjamins, pp. 87-100.
- Künzli, A. & Moser-Mercer, B. (1995): "Human strategies for translation and interpretation", in L. Dreschler-Fischer & S. Pribbenow (eds.), *KI-95 Activities: Workshops, Posters, Demos*, Bonn, Gesellschaft für Informatik, pp. 304-306.
- Kurz, I. (1989): "Conference interpreting: User expectations", *ATA Proceedings of the 30th Annual Conference*, Medford, NJ, Learned Information Inc, pp. 143-148.
- Kurz, I. (1993): "Conference interpretation: Expectations of different user groups", *The Interpreter's Newsletter*, 5, pp. 13-21.
- Massaro, D.W. (1989): *Experimental Psychology: An Information Processing Approach*, San Diego, Harcourt Brace Jovanovich Publishers.
- Meyer, C.F.; Morris, R.A. & Blachman, E. (1994): "Can you see whose speech is overlapping?", *Visible Language*, 28 (2), pp. 110-133.
- Moser, P. (1995): *Simultanes Konferenzdolmetschen. Anforderungen und Erwartungen der Benutzer*, Geneva, AIIC.
- Moser-Mercer, B.; Künzli, A. & Korac, M. (1996): *Physiological and Psychological Stress in Conference Interpreters: Effect on Quality of Output*, (Working paper), Geneva, Ecole de Traduction et d'Interprétation.
- Pöchacker, F. (1994): *Simultandolmetschen als komplexes Handeln*, Tübingen, Gunter Narr.
- Rehbein, J.; Kameyama, S. & Maleck, I. (1994): "Das reziproke Muster der Terminabsprache. Zur Modularität von Diskursen und Dialogen", *Verbmobil memo 23*, Hamburg, Universität Hamburg.
- Rogers, C.R. (1961): *On Becoming a Person: A Therapist's View of Psychotherapy*, London, Constable.