



The International Journal for
Translation & Interpreting
Research

trans-int.org

On the applicability of Internet-mediated research methods to investigate translators' cognitive behaviour

Christopher D. Mellinger

Walsh University

cmellinger@walsh.edu

DOI: ti.106201.2015.a03

Abstract: Experimental research into cognitive processes in translation is often faulted for small sample sizes and the lack of generalisability of findings. Researchers face constraints when identifying and recruiting participants for studies, since the traditional site-bound laboratory imposes geographic limitations on who can potentially participate. Internet-mediated research is a potential solution to this issue, as it expands the size of participant pools and eliminates the need for participant travel. Moreover, different participant variables can be introduced into research projects since scholars are not confined to the local testing area. Keystroke logging in particular is one data collection method that is well suited for Internet-mediated research, yet has not been widely employed in this manner. Here, the literature is reviewed as it relates to Internet-mediated research, with particular emphasis on keystroke logging and translation process research. Specific considerations must be addressed before researchers undertake this type of research project, including the following: the participants themselves; ethics and confidentiality; ecological validity; data security; hardware and software; and the desired measures. Internet-mediated research should be seen as a complement to laboratory studies rather than their replacement.

Keywords: keystroke logging, online data collection, translation process research, Internet-mediated research

1. Introduction

Translation process research (TPR) is an umbrella term that encompasses several phenomena, including the cognitive behaviour and processes of translators as well as the transformation of text or content within the multilingual document lifecycle (Malmkjær 2000; Muñoz Martín 2010). The study of both areas has increased substantially in the past few decades, and scholars continue to draw on research from several disciplines to better understand the translation process. The interdisciplinary nature of TPR is particularly evident in the study of translators' cognitive processes and behaviour (e.g. Tirkkonen-Condit & Jääskeläinen 2000; Shreve & Angelone 2010). Findings from psychology, cognitive linguistics, expertise studies, and other related disciplines have served as the foundation for cognitive research that positions translation as the object of investigation.

Experimental studies, in particular, benefit from research methodologies and protocols that have been developed in the social sciences. For example, translation studies scholars use think-aloud protocols (TAPs), retrospective verbalizations,

and guided interviews to indirectly observe translator behaviour.¹ More recently, researchers have used keystroke logging and eye-tracking to great effect with regard to the analysis of real-time cognitive processes. Physiological metrics, such as pupilometrics, have also been successfully integrated into empirical studies to triangulate elicited data with other qualitative and quantitative measures.²

In order to experimentally investigate the cognitive behaviour of the translator, researchers engage translators to take part in studies designed to elicit process data. Many of these studies are limited in scope, given the considerable challenge of identifying, recruiting, and testing substantial numbers of translators. Such limits are compounded by the geographic restraints imposed by experimental research, since researchers must often study translator behaviour in a controlled, laboratory setting with special equipment and software.

Some researchers have addressed the issue of site-bound laboratory research by including translation students in their studies. This research has provided considerable insight into the differences between novices and experts and the development of expertise. Other scholars (e.g. O'Brien 2006; Dam & Zethsen 2012) have tested groups of translators that work in specific agencies or organizations in an attempt to expand participant pools. Nevertheless, the commonality running throughout many of these translation process and cognitive behaviour studies is the limited size of participant pools. Small sample sizes in experimental research that involve translators are understandable and oftentimes unavoidable, particularly in light of the various permutations of participant variables that one might wish to control — e.g. language combination and directionality, years of experience, native language, and area of expertise.

Several challenges are inherent in studies that include relatively small sample sizes. One illustrative example is O'Brien's (2008) eye-tracking study on how fuzzy matches are processed when a translator uses a translation memory tool. O'Brien's (2008) findings suggest a relationship between processing speed and the match value of the fuzzy match. Likewise, the author demonstrates that translators often compare the source and target segments for differences and do not use the fuzzy match value that is presented with the source and target segment pair. The total number of participants that were included in the study was five. Nevertheless, the author notes:

[...] the possibility that the results might change if the number of participants were increased tenfold. However, given the time-consuming nature of this type of research and the difficulties in acquiring appropriate and equally competent participants who can touch-type, we will have to make do with a small sample size for the present. (p. 97)

These comments highlight the difficulty in the identification and recruitment of participants for a study that requires a specific skill set. Moreover, the author's acknowledgement of a potential change in results with a larger sample size indicates the challenge to generalize results from a small number of participants.

In light of the challenges presented by studies that involve small sample sizes and by the geographic constraints inherent in site-bound laboratory studies, the

¹ Jääskeläinen (2002) provides an extensive review of over one hundred studies that use TAPs. Several more recent examples of retrospective verbalizations or guided interviews include Englund Dimitrova & Tiselius (2014) and Massey & Ehrensberger-Dow (2011).

² It should be noted that many studies now use several of these methodologies to provide multiple perspectives of the object of investigation. A number of scholars have addressed this in the literature, including: Alves et al. (2009); Angelone (2010); Hansen (2010); Shreve & Angelone (2010), among others.

question arises as to how to adequately address this issue. One potential solution, and the argument presented here, is the use of an Internet-mediated research methodology to expand participant pools. The focus of this article is to review the literature on Internet-mediated research. Specifically, this paper proposes that Internet-mediated research methods are a viable option for data collection, and can elicit data often used to investigate cognitive processes in translation. Particular emphasis is placed on keystroke logging, and as such, the paper also reviews several considerations that researchers must address when trying to use keystroke loggers to collect process data from participants via the Internet. While not exhaustive, these areas ought to be considered in the design of a research project. Finally, some conclusions are drawn about Internet-mediated research methodologies and their place in translation studies.

2. Internet-mediated Research

Hewson et al. (2003, p. 1) define Internet-mediated³ (or Internet-based) research simply as “conducting research on the Internet.” The authors classify this type of research as either *primary* or *secondary*, with the former relating to data elicitation from participants, while the latter referring to secondary information sources. For the purposes of this article, Internet-mediated research will refer to Hewson et al.’s (2003) notion of primary research; participant-provided data is necessary to investigate cognitive behaviour, and thus is most appropriate in this context.

Translation studies scholars have adopted several research protocols to conduct primary research via the Internet. An oft-cited methodology to elicit data from participants is the use of questionnaires or surveys. This methodology has been used with relative success to elicit data from respondents on a variety of topics, such as translator compensation (e.g. DePalma & Stewart, 2012) and translation tool usage (e.g. Lommel, 2004). Internet-based questionnaires, though, may not be appropriate for examining the cognitive processes of translators, since introspection about one’s own translation process is nigh impossible. Muñoz Martín (2010, p. 180) asserts its problematic nature, stating: “cognitive psychology rejects introspection, and cognitive philosophy also casts doubt on it.”

While surveys and questionnaires continue to be the research protocols that are used more regularly to collect data over the Internet in translation studies, other research protocols have also been used. One example is Christensen & Schjoldager’s (2011) investigation of translation memory (TM) technology and its impact on cognitive processes. The study relies on participants’ retrospective comments to identify how the translation task is altered by the use of a TM. Respondent comments suggest that the translation with translation memory task differs largely in the drafting phase of text production when compared to translation from scratch. The authors are quick to observe the possible criticism that the comments “may not be consistent with what actually goes on in the subject’s minds, mainly because of the unavoidable delay between the actual processes and the verbalizations” (p. 122). This critique is not specific to the online modality of their data collection methodology, and is often the same criticism levelled against retrospective verbalizations conducted in person. As such, the authors’ innovation of eliciting free-response data via the Internet about

³ Saldanha & O’Brien (2013) use the term *Internet-mediated research*, while Hewson et al. (2003) use the term interchangeably with *Internet-based research*. The term *Internet-mediated research* will be adopted in this article for the sake of consistency.

the translation process merits greater attention and could prove fruitful in future investigations.

The largely qualitative nature of retrospective verbalizations and surveys has led researchers to develop additional measures and propose models to better understand the cognitive processes of translators.⁴ Krings's (2001) tripartite model of cognitive effort in post-editing of machine translation in particular has led to the development of new metrics to measure cognitive processes in translation. In this model, Krings classifies the type of effort required of translators into three categories: technical effort, temporal effort, and cognitive effort. Technical effort encompasses the physical manipulation of hardware, while temporal effort refers to the amount of time required to complete a given task. Cognitive effort, then, is the mental processing required of the translator. The type of effort described is not mutually exclusive, and as Lacruz et al. (2012, p. 1) attest: "in [Krings's] view, temporal effort results from a combination of cognitive and technical effort. Temporal and technical effort can be measured accurately with the help of modern technology."

In line with Lacruz et al.'s (2012) assertion, scholars have adopted additional methodologies in translation process research to measure cognitive effort. One particularly well suited to measuring technical effort is keystroke logging. Keystroke logging, as the name implies, aims to "monitor a user's keyboard actions" (Sagiroglu & Canbek 2009). While keyloggers at a minimum record the user's physical manipulation of a computer keyboard, more sophisticated recording software can "track virtually anything running on a computer" (ibid). The ability to track most, if not all, of a user's behaviour on a computer can be considered both a strength and a weakness, and these considerations are discussed later in this paper; however, the benefit of recording the technical effort exerted by a translator is particularly significant in translation studies. Jakobsen (1999, p. 12) describes the benefit of this type of observation: "Instead of seeing only the final product [...] we can observe all the underlying, preliminary layers of text and decision-making that contributed to the making of the final version." Leijten & Waes (2013, p. 360) echo these remarks, stating that "writing fluency and flow reveals traces of the underlying cognitive processes." Since researchers can gain access to real-time decision-making and problem-solving behaviour via keystroke logs, previously inaccessible cognitive behaviour now can be subjected to greater scrutiny.

Technical effort is not the only type of measure that needs to be recorded, though; temporal information is equally important. Keystroke loggers are one way to obtain this type of data, but other options are available. Such is the case of web-based translation memory systems, machine translation systems, and post-editing systems. For example, Guerberof (2009) uses a post-editing tool developed by Crosslang to capture post-editing behaviour and the associated temporal information. In this study, the researcher investigated how quickly participants processed segments that were either from a translation memory or from a machine translation system. Guerberof's findings differentiate between temporal measures (e.g. time on task and time in segment) and quality measures (e.g. number of errors). Similarly, Denkowski et al. (2014) use a combination of a keystroke logger and machine translation system to evaluate temporal measures associated with cognitive effort when translators post-edit machine translation output. This

⁴ The researcher acknowledges that both retrospective verbalizations and questionnaires can include quantitative data (e.g. coding free response comments or the inclusion of Likert-type scale measures); however, these measures do not provide enough cognitive process data to be used independently for a complete view of the translator's cognitive behaviour.

approach to data collection differentiates between temporal measures (e.g. time per segment) and technical measures (e.g. number of mouseclicks or keystrokes), which ultimately provides a more comprehensive overview of the post-editor's behaviour.

The question arises as to how keystroke logging or other software can be implemented in Internet-mediated research. One of the most often used pieces of keystroke logging software in translation studies is *Translog*. Developed by Jakobsen (1999), this software allows for scholars to test participants in a laboratory setting to record translation behaviour. Unfortunately, a server-based version of *Translog* is not available, and thus cannot be hosted on a web server to allow for Internet-mediated research. Consequently, translation scholars have needed to look for other keyloggers to conduct research online.

Hewson et al. (2003, p. 47) argue that “essentially, almost any piece of research that could be implemented offline using a computer program can also be implemented online over the Internet.” Indeed, several researchers have successfully integrated keystroke logging into online research studies and have collected cognitive process data. For example, Lacruz, et al. (2012) collected data using *TransCenter* to investigate cognitive effort in post-editing. Likewise, Mellinger (2014) employed *TransCenter* to examine participants' cognitive effort when working with translations proposed by a translation memory or when translating segments without any aid. Initially developed by Denkowski & Lavie (2012) for research on post-editing of machine translation, this tool allows researchers to present a source and target text to participants and record their translation or editing behaviour. To borrow Krings's terminology, the technical effort is recorded, as are measures indicative of temporal effort.

The several examples of online data collection cited thus far demonstrate Internet-mediated research is a viable means of investigating cognitive processes in translation. Further research is necessary to determine the comparability of results between the online and on-site modalities of participant behaviour; however, research on web usability studies (e.g. Tullis et al., 2002) indicates very similar results in participant performance regardless of whether the study was conducted via the Internet or in the laboratory. Likewise, the adoption of online process data collection in related areas of inquiry, such as digital writing (e.g. Waes et al., 2012), post-editing (e.g. Roturier et al., 2013), and psychology (e.g. Reips, 2002), further supports its inclusion in translation process research.

3. Considerations for Internet-mediated Research

With the foundation laid for Internet-mediated research in translation studies, researchers are tasked with adapting methods used in laboratory settings to accommodate the specific requirements of online data collection. Keystroke logging programs are no exception, despite being particularly well suited for online use. There are several considerations that must be addressed, and each will be reviewed in turn here. Specifically, six topics are addressed below, as are their strengths and weaknesses. They are: (1) participants; (2) ethics and human subjects; (3) ecological validity; (4) data security; (5) hardware and software; and (6) measures. The topics are discussed with full awareness that these are considerations that are more practical and cannot be considered an exhaustive review of all the issues arising in during Internet-mediated research. Instead, the aim here is to highlight some of the challenges faced by researchers when conducting Internet-mediated research.

3.1 Participants

One of the main benefits of Internet-mediated research and online data collection methods is the size of the participant pool. Research that is conducted in a laboratory setting places geographic constraints on who can participate. In translation studies, this issue is further compounded by the skill set that participants need to possess (namely, being translators). Dumas & Fox (2012, p. 1227) note the benefits of testing participants remotely in usability studies:

- “You can reach a worldwide population of participants because you are not limited to the local testing area. This may be especially helpful when there are not many users, and they are geographically dispersed.
- It is easier to get a participant to volunteer because they do not have to travel. [...]
- You do not need a usability lab.”

While these comments are specific to another field of study, the benefits are directly applicable to translation process research. The notion of geographically-disperse participants is of particular import to translation studies, as there may not be a large number of translators located within travelling distance to a university. As noted previously, this issue is compounded when researchers are interested in investigating translators with specific characteristics — e.g. years of experience, language combination and directionality. Translation process research often has worked with populations of convenience, such as translation students, and with language combinations that are more readily accessible. Internet-mediated research extends the reach and scope of potential projects, since it allows scholars to recruit participants that reside or work far from any lab setting. Moreover, research with distinct populations allows researchers to corroborate findings on cognitive processes with different samples.

The convenience factor of not requiring participant travel is equally beneficial. This line of reasoning is similar to the previously presented argument for expanded participant pools, in that translators with specific characteristics of interest may be more willing to participate if they are not required to travel extensively. Birnbaum (2004) notes the challenges of participant recruitment, and provides an extensive review of recruitment practices when conducting Internet-mediated research. In particular, the author notes lower response rates from unsolicited e-mail messages. The depersonalized nature of Internet-mediated research could be responsible for this challenge. Consequently, TPR scholars must develop a recruitment plan that addresses this issue. Researchers may wish to adopt a multi-pronged approach when contacting potential participants; recruitment scripts and messages can be circulated using e-mail, professional listservs and blogs, and personal contacts in an effort to increase response rates.

One final comment on studying participants using Internet-based data collection is whether the research is conducted in a synchronous or asynchronous manner. Dumas & Fox (2012) differentiate the two based on the involvement of the researcher at the time of the study. Synchronous testing requires the researcher to directly interact with the user when effecting the study, whereas asynchronous testing does not. Dumas & Fox (2012) note that larger research pools can be reached when remote testing is done asynchronously, as time requirements of the researcher are not a limiting factor to the number of participants who take part in the study. In short, Benfield & Szlemko (2006, n.p.) succinctly synthesize these benefits: “if utilized properly, [Internet-based data collection] can reduce costs and make unfunded projects feasible, yield larger and more representative samples, and obviate hundreds of hours of data entry.”

3.2 Ethics and Human Subjects

As with any research with human subjects, scholars must take care to ensure the confidentiality of participants, eliminate conflicts of interest, obtain informed consent, and minimize potential risks. Guidelines on this type of research are provided by Human Subjects and Institutional Review Board of research institutions, and should be followed when planning and conducting Internet-mediated research. The implementation of these guidelines is unique to each research project, and a discussion of all the potential permutations lies outside the scope of this article. Nevertheless, brief mention should be made to confidentiality as it differs slightly from site-bound research projects.

One of the conditions that must be followed when conducting research of any kind is the maintenance of the confidentiality of personally-identifiable information. Site-bound research can mitigate most of this risk by providing a private space to conduct the study, as well as by assigning a unique ID number for each participant that cannot be directly associated with the person. In contrast, research conducted via the Internet poses a unique challenge, in that the IP address of a participant can be recorded and stored. This information could be used to identify the participant, which in turn would link the information provided in the study to the specific person. Benfield & Szlemko (2006) suggest removing this information from any data set early, or when possible, electing not to record the IP address. Researchers must make this determination prior to the study in order to mitigate any possible risk of linking the participant with a specific set of data. This possibility is not a limitation per se, but researchers should be mindful of the potential drawbacks it presents when designing any process research that will be conducted via the Internet.

3.3 Ecological Validity

One of the perennial debates that often arises during the evaluation of translation process research is whether the results obtained in a laboratory setting are generalisable to actual working conditions of professional translators. This question is important to consider, given the complex nature of the translation process. A benefit to Internet-mediated research is that translators can participate in the study using their own hardware and software, and in their normal working environment (Dumas & Fox 2012). This obviates the need for participants to become acclimated to the physical constraints of a laboratory setting, while still controlling a number of variables in the study itself.

The control of experimental conditions presented via the Internet, though, introduces other concerns for ecological validity. Many of the software packages that have been cited thus far are not commercially available tools with which translators regularly work. Moreover, these systems necessarily restrict translator behaviour to ensure that the data elicited in the study correspond with the dependent variables the researcher wants to observe. Software restrictivity will be illustrated by means of TransCenter (Denkowski & Lavie 2012) as an example. TransCenter is designed to elicit process data on post-editing, and does so via keystroke logging and a 5-point Likert-type scale. Research participants are asked to work on a segmented text and also asked to evaluate each segment that has been post-edited as they move through the task. One issue that could be raised as to the ecological validity of the task is that evaluation, albeit implicitly performed by post-editors, is not explicitly requested during commercial post-editing projects. Likewise, TransCenter is not a post-editing tool that is used commercially, thus introducing some unfamiliarity into the task on the part of research participants. These concerns are similar to those levelled against laboratory research conducted on-site, and similar measures must be taken to

compensate. For example, a practice round of the task might help familiarize the participant with the tool; written or online documentation for the software might also benefit the translator, as would upfront training to use the system.

3.4 Data Security

One of the challenges when conducting Internet-mediated research is preserving data and maintaining its security during the transmission between participants and the researcher. Benfield & Szlemko (2006, n.p.) note this difficulty, and highlight that “data are most susceptible to hacking, corruption, etc., while these are being transferred from the respondents’ computers to the researchers’ computer.” Certainly, the corruption of data during its transit from participant to researcher is troublesome, since any loss or change may alter the results of the experiment. One possible rationale for corrupted data may be unreliable Internet connections; however, as Birnbaum (2004) suggests, most people with Internet access at the time of his writing will have a connection that is suitable for research purposes. In addition, data encryption is important and recommended to mitigate the potential for third-parties to intercept the transmitted data. As an example, many online survey software packages (e.g. Qualtrics or SurveyMonkey) incorporate the option to encrypt data during the collection and transmission phase. Researchers would be wise to enable these features in order to mitigate some of the risks associated with online data transmission. An in-depth discussion of the specifics of how each of these features are enabled is outside the scope of this paper — given the ever-expanding range of software packages that researchers may potentially adopt — however, many of these options are relatively user-friendly and may be addressed in the software documentation.

The transmission stage, however, is not the only point in the collection process in which data are susceptible; data storage must also be considered as a potential vulnerability (Benfield & Szlemko, 2006). Just as researchers must account for how data will be stored from laboratory studies, so too must scholars find ways to reasonably secure data obtained from Internet-mediated research. Files can be password-protected and stored using university computing resources, greatly improving the level of protection compared to that provided by an individual’s computer. Brief mention should be made of cloud storage, as services like Dropbox, Google Drive, and SkyDrive make significant inroads into the consumer market. Kaufman (2009) highlights several concerns with cloud storage security, such as data ownership, accessibility, and vulnerability. These issues should be examined with great care, should a scholar decide to use these storage systems. Perhaps the best way to address data security concerns related to cloud storage, though, is to avoid their use altogether when saving research data. This issue merits greater attention; however, it falls outside the scope of this particular article.

The sizeable body of literature relying on keyloggers in translation studies clearly demonstrates their usefulness in research projects, yet Sagioglu & Canbek (2009) note that keyloggers can be used to more malicious ends. Ill-disposed users of keyloggers, for example, can compromise personal information and transmit such information to third-parties. The authors continue this line of reasoning and attest to the potential use of keyloggers in identity theft (*ibid.*, p. 14–15). As a result of this devious use, many antivirus and anti-malware software packages may block their use in browsers. Server-side programming may be one solution to ensure that participants are able to run the experiment in their Internet browsers (Birnbaum, 2004).

3.5 Hardware and Software

Up to this point, the four topics that have been reviewed are more closely related to the design of the research study. In addition to these more theoretical and conceptual considerations, the actual implementation of the study have requirements that are specific to Internet-mediated data collection. As noted, keystroke logging records the interaction a user has with a physical keyboard. Sagioglu & Canbek (2009) describe two main types of keyloggers that can be used to achieve this effect: hardware and software keyloggers. The first type — hardware keyloggers — are physical devices that can be integrated into the keyboard itself. Software keyloggers, in contrast, do not require a physical device and instead collect the data within the operating system of the user. No additional hardware is required, and can still achieve the same amount of data. In the context of Internet-mediated research, researchers may want to employ software keyloggers since they do not require any additional hardware and can be used via the Internet. Scholars may find hardware keyloggers cost-prohibitive if participants need to be provided with specialized equipment to participate in the study.

A limitation of software keystroke loggers is the availability of several different keyloggers that are specific for research purposes and can be used online. Waes et al. (2012) note a number of keyloggers that are available to investigate text production, including Inputlog, ScriptLog, and Translog. The authors provide a comparison of the software's features, which highlights the strengths of each. Other tools that have been used to collect process data in translation process research have been cited previously in this article, such as TransCenter. The challenge for researchers is to identify the tool that is best suited for the research questions that have been posed. Moreover, Internet-mediated research necessarily requires these tools to be deployable in a server environment so as to allow users to access these tools via the Internet. Further development of new tools is crucial to allow researchers to capture process data via the Internet and triangulate several measures to gain greater insight into the translator's cognitive processes.

Beyond the hardware and software specific for data collection, brief mention should be made to the participant's hardware and software. Given that each of the participants will ostensibly work at their own computer workstation, significant variation in computer configuration is expected to occur. Reips (2002, p. 246) cautions researchers who conduct Internet-mediated research not to discount the technical variance in “web browsers, net connections, hardware computers, and so forth.” To counteract the real possibility of variations impacting the results of the study, researchers are advised to test the experiment under a variety of conditions. Reips notes several conditions that should be tested, such as different web browsers, operating systems, and internet connection speeds. To this list could be added antivirus systems, browser plug-ins, and character set configurations. Obviously, a good faith effort is required as scholars cannot possibly test every possibility. If a specific combination or set of conditions are preferred, researchers can include the necessary settings in the experiment instructions to participants.

3.6 Measures

To examine cognitive processes, studies have relied on technical and temporal measures to gauge the underlying construct of cognitive effort. In particular, pauses can elucidate cognitive processing in participants (Butterworth, 1980; Lacruz & Shreve, 2014; Leijten & Waes, 2013). Pause metrics can be obtained from keystroke logging, and tools that facilitate the extraction of pause data would be highly desirable in such studies. Leijten & Waes (2013, p. 360) cite the “analytical focus on pause (length, number, distribution, location, etc.) and

revision (number, type, operation, embeddedness, location, etc.) characteristics” when investigating cognitive processes in writing. Thus, keystroke logging data can be particularly desirable to collect, as this data reveals several different measures.

Likewise, retrospective comments and questionnaires can be used to elicit participant data, which can be triangulated with the aforementioned measures. As Christensen & Schjoldager (2011) demonstrate, retrospective verbalizations can be collected in an online modality. Similarly, questionnaire data has been shown to be comparable in on- and off-line settings (Stanton, 1998). Both of these measures, while not specifically eliciting cognitive process data, can support research on translator behaviour and better elucidate the translation process.

4. Conclusion

In light of the recurrent challenges in translation process research involving small sample sizes, researchers are tasked with finding innovative ways to identify and recruit participants to take part in their studies. The traditional site-bound laboratory setting compounds the issue by imposing geographic constraints on the potential participants who could be included. Internet-mediated research is one clear solution to expanding the size of these participant pools. Moreover, research conducted via the Internet ostensibly extends the scope of potential research projects by allowing researchers to investigate cognitive processes in participants who have distinct characteristics. Thus far, scholars have only begun to use Internet-mediated research to investigate the translation process. These successful initial attempts at data collection online using keystroke logging software demonstrate the viability of this approach to research, and additional experiments are needed to verify its use.

Nevertheless, scholars cannot simply abandon all other data collection methods, and instead must be mindful of several practical considerations when designing research projects that will be conducted via the Internet. Moreover, the specific parameters imposed by Internet-mediated data collection must be addressed when designing the study. These considerations include participant selection, ethical research practices, and the ecological validity of the proposed experiment. Researchers must also account for data security, acquire hardware and software that can accommodate this type of data collection, and ensure that the measures being obtained are appropriately contextualized.

Lastly, it should be expressly noted that the argument here is not to replace laboratory studies with Internet-mediated research. Instead, these two modalities should be seen as complementary, allowing cognitive processes data to be collected from a greater number and wider variety of participants. Internet-mediated research has its limitations as have been outlined in the previous sections, yet also shows its value. Further methodological research is necessary to develop research protocols that can incorporate online data collection into translation process research.

About the author

Christopher D. Mellinger is Assistant Professor of Spanish at Walsh University, where he teaches medical translation and interpreting as well as courses on Spanish for healthcare. Mellinger holds a Ph.D. in Translation Studies and a M.A. in Translation (Spanish) from Kent State University. Mellinger is an ex-officio board member of the American Translation and Interpreting Studies Association and the managing editor of the journal *Translation and Interpreting Studies*. His research interests include translation and cognition, translation process research, and translation technology.

References

- Alves, F., Pagano, A., & da Silva, I. (2009). A New Window on Translators' Cognitive Activity: Methodological Issues in the Combined Use of Eye Tracking, Key Logging, and Retrospective Protocols. In I. Mees, F. Alves, & S. Göpferich (Eds.), *Methodology, Technology and Innovation in Translation Process Research* (pp. 267–291). Copenhagen: Samfundslitteratur.
- Angelone, E. (2010). Uncertainty, Uncertainty Management, and Metacognitive Problem Solving in the Translation Task. In G. M. Shreve, & E. Angelone (Eds.), *Translation and Cognition* (pp. 17–40). Philadelphia: John Benjamins Publishing Company.
- Benfield, J. A., & Szlemko, W. J. (2006). Internet-Based Data Collection: Promises and Realities. *Journal of Research Practice*, 2(2), n.p.
- Birnbaum, M. H. (2004). Human Research and Data Collection via the Internet. *Annual Review of Psychology*, 55, 803–832.
- Butterworth, B. (1980). Evidence from Pauses in Speech. In B. Butterworth (Ed.), *Language Production: Speech and Talk* (Vol. 1, pp. 155–176). London: Academic Press.
- Christensen, T. P., & Schjoldager, A. (2011). The Impact of Translation-Memory (TM) Technology on Cognitive Processes: Student-Translators' Retrospective Comments in an Online Questionnaire. In B. Sharpe, M. Zock, M. Carl, & A. L. Jakobsen (Eds.), *Proceedings of the 8th International NLPCS Workshop: Special theme: Human-Machine Interaction in Translation* (pp. 119–130). Copenhagen: Samfundslitteratur.
- Dam, H. V., & Zethsen, K. (2012). Translators in international organizations: A special breed of high-status professionals? Danish EU translators as a case in point. *Translation and Interpreting Studies*, 7(2), 212–233.
- Denkowski, M., & Lavie, A. (2012). TransCenter: Web-Based Translation Research Suite. *AMTA 2012 Workshop on Post-Editing Technology and Practice Demo Session*.
- Denkowski, M., Lavie, A., Lacruz, I., & Dyer, C. (2014). Real Time Adaptive Machine Translation for Post-Editing with cdec and TransCenter. *Proceedings of the EACL 2014 Workshop on Humans and Computer-assisted Translation*, (n.p.).
- DePalma, D. A., & Stewart, R. G. (September 2012). *Trends in Translation Pricing: Falling Rates Accompany Changes in Economy and Buying Behaviors*. Lowell, MA: Common Sense Advisory.
- Dumas, J. S., & Fox, J. E. (2012). Usability Testing. In J. A. Jacko (Ed.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications* (pp. 1221–1241). New York: CRC Press.
- Englund Dimitrova, B., & Tiselius, E. (2014). Retrospection in Interpreting and Translation: Explaining the Process? In R. Muñoz Martín (Ed.), *Minding Translation/Con la traducción en mente*, Special issue of *MonTI*, 1(1), 177–200.
- Guerberof, A. (2009). Productivity and Quality in the Post-Editing of Outputs from Translation Memories and Machine Translation. *The International Journal of Localisation*, 7(1), 11–21.

- Hansen, G. (2010). Integrative Description of Translation Processes. In G. M. Shreve, & E. Angelone (Eds.), *Translation and Cognition* (pp. 189–212). Philadelphia: John Benjamins Publishing Company.
- Hewson, C., Yule, P., Laurent, D., & Vogel, C. (2003). *Internet Research Methods: A Practical Guide for the Social and Behavioural Sciences*. London: SAGE Publications.
- Jääskeläinen, R. (2002). Think-aloud Protocol Studies into Translation: An Annotated Bibliography. *Target*, 14(1), 107–136.
- Jakobsen, A. L. (1999). Logging Target Text Production with Translog. In G. Hansen (Ed.), *Probing the Process in Translation: Methods and Results* (pp. 9–20). Copenhagen: Samfundslitteratur.
- Kaufman, L. M. (2009). Data security in the world of cloud computing. *Security & Privacy, IEEE*, 7(4), 61–64.
- Krings, H. P. (2001). *Repairing Texts: Empirical Investigations of Machine Translation Post-editing Processes*. (G. S. Koby, Ed.) Kent, OH: Kent State University Press.
- Lacruz, I., & Shreve, G. M. (2014). Pauses and Cognitive Effort in Post-Editing. In S. O'Brien, L. W. Balling, M. Carl, M. Simard, & L. Specia (Eds.), *Post-Editing of Machine Translation: Processes and Applications* (pp. 244–272). Cambridge: Cambridge Scholars Publishing.
- Lacruz, I., Shreve, G. M., & Angelone, E. (2012). Average Pause Ratio as an Indicator of Cognitive Effort in Post-Editing: A Case Study. *Proceedings of the AMTA 2012 Workshop on Post-editing Technology and Practice*. San Diego, CA.
- Leijten, M., & Waes, L. v. (2013). Keystroke Logging in Writing Research: Using Inputlog to Analyze and Visualize Writing Processes. *Written Communication*, 30(3), 358–392.
- Lommel, A. (2004). *Translation Memory Survey: Translation Memory and Translation Memory Standards*. Romainmôtier, Switzerland: LISA. Accessed 13 March 2014. Retrieved from <http://bit.ly/PyJvKH>
- Malmkjær, K. (2000). Multidisciplinarity in Process Research. In S. Tirkkonen-Condit and R. Jääskeläinen (Eds.), *Tapping and Mapping the Processes of Translation and Interpreting: Outlooks on Empirical Research* (pp. 163–164). Philadelphia: John Benjamins Publishing Company.
- Massey, G., & Ehrensberger-Dow, M. (2011). Commenting on Translation: Implications for Translator Training. *The Journal of Specialised Translation*, 16, 26–41.
- Mellinger, C. D. (2014). Computer-assisted Translation: An Empirical Investigation of Cognitive Effort. (Unpublished Ph.D. dissertation, Kent State University, Kent, OH). Retrieved from <http://bit.ly/1ybBY7W>
- Muñoz Martín, R. (2010). On Paradigms and Cognitive Translatology. In G. M. Shreve, & E. Angelone (Eds.), *Translation and Cognition* (pp. 169–187). Philadelphia: John Benjamins Publishing Company.
- O'Brien, S. (2008). Processing Fuzzy Matches in Translation Memory Tools: An Eye-tracking Analysis. In S. Göpferich, A. L. Jakobsen, & I. Mees (Eds.), *Looking at Eyes: Eye-Tracking Studies of Reading and Translation Processing* (pp. 79–102). Frederiksberg: Samfundslitteratur.
- O'Brien, S. (2006). Eye-Tracking and Translation Memory Matches. *Perspectives: Studies in Translatology*, 14(3), 185–205.
- Reips, U.-D. (2002). Internet-Based Psychological Experimenting: Five Dos and Five Don'ts. *Social Science Computer Review*, 20(3), 241–249.
- Roturier, J., Mitchell, L., & Silva, D. (2013). The ACCEPT Post-Editing Environment: a Flexible and Customisable Online Tool to Perform and Analyse Machine Translation Post-Editing. In S. O'Brien, M. Simard, & L. Specia (Eds.), *Proceedings of MT Summit XIV Workshop on Post-editing Technology and Practice* (pp. 119–128). Nice, France.
- Sagioglu, S., & Canbek, G. (2009). Keyloggers: Increasing Threats to Computer Security and Privacy. *IEEE Technology and Society Magazine*, 11–17.
- Saldanha, G., & O'Brien, S. (2013). *Research Methodologies in Translation Studies*. Manchester: St. Jerome Publishing.
- Shreve, G. M., & Angelone, E. (Eds.). (2010). *Translation and Cognition*. Philadelphia: John Benjamins Publishing Company.

- Stanton, J. M. (1998). An Empirical Assessment of Data Collection Using the Internet. *Personnel Psychology*, 51(3), 709–725.
- Tirkkonen-Condit, S., & Jääskeläinen, R. (Eds.) (2000). *Tapping and Mapping the Processes of Translation and Interpreting*. Amsterdam: John Benjamins Publishing Company.
- Tullis, T., Fleischman, S., McNulty, M., Cianchette, C., & Bergel, M. (2002). An Empirical Comparison of Lab and Remote Usability Testing of Websites. *Proceedings from Usability Professional Association Conference* (n.p.). Orlando, FL.
- Waes, L. v., Leijten, M., Wengelin, A., & Lindgren, E. (2012). Logging Tools to Study Digital Writing Processes. In V. W. Berninger (Ed.), *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology* (pp. 507–533). New York: Psychology Press.