A quest for web search optimisation: an evidence-based approach to trainee translators' behaviour

Abstract:

Many translation textbooks consist of chapters dedicated to web resources and web search in general. However, surprisingly little is known about how (trainee) translators actually conduct web search. The present study, therefore, adopts an evidence-based approach to empirically examine how the web is used by trainee translators particularly for terminology purposes. Unlike previous studies that focus on translators' idiosyncrasies and styles, this study taps into the idea of web search optimisation. In other words, it aims to depict what characterises a more optimal web search for translation and asks about what (strategic) web search process trainee translators could engage in in order to optimise their search results. 18 postgraduate translation students were asked to translate three texts on three separate occasions. Their translation and web search processes were screen-recorded and partially verbalised. Data suggest that more optimal web search is characterised by an 'explorer's approach' and more profound engagement with 'secondary actions' rather than 'primary actions'.

Keywords: web search; screen recording; translator training; specialised translation; information literacy

1. Introduction

In recent years, information literacy has been recognised as an important instrumental competence by many translation scholars (see EMT, 2017, p. 9; Göpferich, 2009, p. 12; Massey and Ehrensberger-Dow, 2011; PACTE, 2011, pp. 30-53; Pinto and Sales, 2007, 2008a, 2008b; Sales et al., 2018). Some studies have been carried out on translators' information behaviour in general. For instance, White et al (2008) conducted focus group interviews of 19 professional translators, exploring the nature of their information needs, what types of resources they use and potential information challenges faced by translators working in less commonly used languages. Building on their studies conducted a decade earlier (see Pinto and Sales, 2007, 2008a, 2008b), Sales et al (2018, p. 191) described how 1st-year undergraduate translation students learned and conducted their 'documentation process'. Their studies confirmed that students' predominant and principal source of information was the Internet. The prominent use of the Web as an information source was not just found among trainee translators but also applied equally to professionals. In fact, it has been suggested that freelance translators spent around a third of their time on web-based research activities during their translation process (Gough, 2015, p. 244). Interestingly, according to Massey and Ehrensberger-Dow (2011, p. 205), translators' information behaviour largely correlated with their length and types of experiences, but almost unexpectedly, recent graduates were more efficient and effective than veteran translators in locating relevant information. Therefore, they concluded that the key to information literacy probably lies in training rather than professional experience alone.

1

1.1 Aims

Despite its obvious significance, relatively little research has been dedicated to trainee translators' web search process specifically. Therefore, **this study aims to investigate web search process conducted by trainee translators, when facing terminology problems**. In particular, the following aspects will be explored.

(1) Characteristics of successful vs unsuccessful web search episodes[1] and (potentially) why unsuccessful web search episodes occur

(2) Characteristics of more optimal web search episodes

(3) Strategic web search process a trainee translator may engage in in order to achieve web search optimisation.

In this study, optimisation is deemed to be a balanced combination of efficiency and effectiveness. Efficiency is defined as a minimum amount of time being spent on web search episodes. Effectiveness is concerned with the precision of TL equivalents. If no satisfactory TL equivalent is found for a ST term, a web search episode is deemed unsuccessful. In terms of who deems a TL equivalent satisfactory, the students' points of views is largely taken into account, as their verbalisation normally provides good clues about this. Even after a lengthy search when a student decides to write down something in desperation, his/her verbalisation usually reflects whether a satisfactory TL equivalent is found or not.

The idea of efficiency is a relative concept in this study. This is because inevitably idiosyncrasy exists between individual translators and terminological problems. As a result it will be impossible to pinpoint how long or how many steps constitute an efficient web search. Nevertheless, when comparing different translators' web search episodes for the same ST term, distinction can be made between more efficient episodes and less efficient ones. Regarding the relationship between efficiency and effectiveness, it has to be acknowledged that in practice, efficiency should never precede effectiveness since it would be pointless to aim for efficiency at the expense of effectiveness in translation.

2. Previous studies

This section consists of two parts. First of all, existing scholarly work dedicated to web search behaviour in translation studies will be reviewed. Then, a number of information retrieval (IR) studies will be discussed, particularly those of relevance to the present study with regard to their findings and metrics.

2.1 Web search behaviour in translation

---

[1] A web search episode is defined as a web search process that is associated with a particular terminology. More details about this can be found in Section 3.

2

Enríquez-Raído (2011, 2014) was one of the first scholars who studied the web search process per se. As is the case with the present study, she focused on trainee translators' web search process. Her subject were four undergraduate students in her Spanish/English translation module (despite having two professionals in her pilot study). They were asked to translate two popular scientific texts in week three and week four respectively in the module lasting 12 weeks. Her study can be seen as a multi-subject case study in which she employed multiple data collection methods investigating each of the trainees' individual behaviour. Her data collection methods included background questionnaires, screen recording, online search reports (a type of self-reflective learning report about web search) and semi-structured interviews. Apart from demonstrating the feasibility and advantage of a combination of various data collection methods for web search investigation, Enríquez-Raído's main contribution lies in its pedagogical implication. For instance, she pointed out that the online search report was particularly useful in raising students' awareness of their translation and web search process. She also observed that in these reports, students tended to focus on how they dealt with problematic technical terms rather than general lexical problems (2014, pp. 184-185). In addition, she advocated that information literacy should be developed for each specialised translation field and closely linked to professional practice (2014, p. 183). She also suggested that pedagogy related to (successful) query construction should not just rely on general knowledge about search engine, but more importantly on the combination of information needs analysis and its consequent selection of key ST terms (ibid). All these findings partly inspire and warrant the present study.

Similar to Enríquez-Raído, XXX also focused on trainee translators' behaviour and its pedagogical implications. Her data derived from six postgraduate Chinese trainee translators' screen recording and think-aloud protocols while translating a semi-specialised scientific text. Her study presented an empirical observation of the variety of queries and web resources employed by Chinese trainee translators, and more importantly, how different combinations of web resources were used by individual trainees. For example, she highlighted an interesting phenomena where a subject used a (Chinese) Statistical Machine Translation (SMT) system exclusively as if it was a dictionary. She also noted on how the ergonomics and dynamics of the web search process was dramatically altered when a trainee adopted a hybrid dictionary and information seeking tool online (i.e. Lingoes). In this case, cognitive efforts commonly spent on locating different types of online resources, and going back and forth between them were dramatically minimised; instead, maximum cognitive effort was spent on assessing and deciphering a sizeable number of search results being presented in one single window in the hybrid tool.

Apart from Enríquez-Raído and XXX, Gough (2015) was also one of the latest scholars who focused on web search behaviour specifically. Unlike the other two studies, Gough's subjects were freelance translators of at least 5 years of experiences working from English into other languages. She used the terms, 'online consultation process' and 'translators' interaction with online resources' instead of 'web search behaviour' to describe her data. As a matter of consistency and to avoid confusion, I will only use the term, 'web search behaviour' throughout this paper. Gough investigated freelance translators' web search behaviour both collectively and individually. Her study involved a two-stage data collection design: initially a large scale web-based survey (N=540) and then a screen-recorded translation task followed by post-task questionnaires (N=16). Collectively, she focused on the nature and quantity of web resources used and also the time (spent) on them as a whole in the translation process. Individually, she focused on different types of resources used, how they were used and the time factors involved in using these resources. Her main contribution was the formulation of two taxonomies, i.e. a

3

Resource and Type User Taxonomy and a Translator Resource Style Taxonomy, by means of which she categorised 5 different characteristics or styles of translators' web search behaviour. 'Time factor' appeared to be an important measurement in Gough's study. In fact, her taxonomy of translators' web search styles was principally based on the measurement of time spent on web searching and volume of web resources used. Her five types of translators' (web search) styles comprised the 'explorer', the 'prolific translator', the 'methodical translator', the 'economical translator' and the 'understated translator'. According to Gough (2015, pp. 229-237), the explorer spent the largest amount of time and effort on web search while the prolific translator also made huge effort in terms of the volume of web resources used but somehow spent less time on them. Being the opposite of the explorer, the economical translator spent the minimal amount of time and effort on web search whereas the understated translator spent extended amounts of time but with relatively low web search volume. Finally, sitting in the middle is the methodical translator who spent a moderate amount of time and effort on web search. Even though translators' idiosyncratic styles are not a focus of the present study, it is interesting to see how or whether traces of such styles can be found in trainee translators' behaviour, particularly in relation to the time factor since it is linked to the notion of efficiency. In Gough's study, there was no distinction between optimal and less optimal (nor successful vs unsuccessful) web search episodes and perhaps more crucially, there was little lateral analysis of web search behaviour prompted by particular terminological issues in the source text.

2.2 Web search behaviour in Information Retrieval Studies

Apart from the existing web search studies in translation, it is important to acknowledge that web search behaviour in general is widely studied within the field of Information Retrieval (IR), particularly in its sub-discipline, user studies. Part of the originality of this article is to deploy and adapt relevant concepts and metrics in IR research in investigating web search conducted as part of the translation process.

In IR research, a substantial amount of effort has been spent on analysing 'big data' (Mayer-Schönberger, Cukier and Hogan, 2013), using search logs mined via search engines or other online information systems. Search logs offer vast amounts of data for quantitative analysis on typical web searchers' behaviour, such as query formulation, query length, number of queries per search session and various clicking behaviour, etc (see Jansen et al., 2000; Silverstein et al., 1999; Spink et al., 2000). After reviewing major studies using search logs, Jansen and Pooch (2001, p. 240) concluded that 'the vast majority of web searchers use approximately two terms in a query, have two queries per session, do not use complex query syntax … use of Boolean operators in web queries is almost nonexistent…'. Despite their obvious advantages, these studies also have their limitations, particularly in terms of not distinguishing web search in different contexts and for different purposes. This is part of the gap the present study aims to fill as it seeks to elucidate web search conducted by trainee translators during their translation process.

In order to depict the characteristics of optimal web search in translation, two types of web search behaviour along with their related metrics, namely, 'primary actions' and 'secondary actions', were adopted from the field of Information Retrieval (White, 2016, pp. 21-37). Primary actions indicate actions conducted while web searchers interact with search engines directly. These actions are often related to queries. Secondary actions are interactions with search engine result pages (or SERPs for short) after queries have been submitted. Primary and

4

secondary actions along with their related metrics are briefly illustrated in Table 1 below.

INSERT TABLE 1 HERE

'Query formulation' indicates how queries are structured or formulated. For instance, does a query consist of more than one term or some form of syntax, such as Boolean symbols or the use of natural language? In IR studies, 'query re-formulation' refers to '[a] system based on a neural network that [automatically] rewrites a query to maximise the number of relevant documents returned.' (Nogueira and Cho, 2017, p. 1) In other words, it refers to the function of automatic query suggestion either during or after a query has been submitted in a search engine. However, in the present study, the notion of 'query re-formulation' is adapted to refer to how queries are re-structured or re-formulated by translators (not by search engine algorithms) throughout the course of a web search episode.

'Query duration' stands for the duration (or the length of time) a searcher spends on a query. This is a metric closely linked to the notion of time or efficiency in the present study. Query duration provides little value in isolation. Nevertheless, in the present study, I am more interested in how query duration may vary among search episodes and over time. For the purpose of comparing query durations between episodes, the number of queries submitted in relation to the duration of a whole search episode is calculated, i.e. 'average query duration'.

'Query abandonment' refers to a searcher not clicking on any of the search results in SERPs after query submission. It is often seen as an indication of low satisfaction or frustration with the search (see Khabsa et al., 2016; Wu and Kelly, 2014). In other words, the searcher is not satisfied with the search results being presented in SERPs and therefore either frequently abandons the query or the whole search episode.

'Dwell time' indicates how long a searcher spends on a webpage or website accessed via SERPs. This is the other metric that is related to the notion of time and efficiency. In IR research, 'dwell time' is often associated with how relevant a retrieved document or webpage is perceived to be by the searcher (White, 2016, p. 34). Some studies even suggest that a dwell time of more than 30 seconds is a good indication that a searcher is somewhat satisfied with the content of the webpage (Kelly and Belkin, 2004).

A trail comprises 'a temporally ordered sequence of [webpages] visited by searchers' (Bilenko and White, 2008). In the present study, the idea of 'trail length' is slightly adapted to indicate how many webpages in SERPs are visited. Interestingly, White and Drucker (2007) identified two types of searchers based on their trails: navigators and explorers. While navigators are more inclined to follow a short and directed trail, explorers tend to backtrack, or hop between different webpages more (White, 2016, p. 30). In other words, a navigator's trail is more linear whereas an explorer's trail tends to be more fuzzy and sometime overlapping. It is worth noting that Gough also used the term, 'the explorer' (2015, p. 229) in her taxonomy. Similar to White, she described 'the explorer' as having a type of research path that branches out more. It will be

5

interesting to see whether the same type of characteristics can be found in the present study.

3. Data collection

In order to investigate web search conducted by trainee translators, a corpus of 18 postgraduate students from the MSc Specialised Translation programme at the researchers' academic institution were asked to translate three English texts (roughly 500 words each) into Chinese (mother tongue) on three separate occasions in the academic years 2015/16 and 2016/17. Of the 18 students, 17 were female and 1 was male, a not uncommon gender ratio of student cohorts in this programme. This study adapts a quasi-naturalistic design where students were asked to screen-record (and verbalise whenever possible) their translation and web search process in their normal place of work (either at home or at an open-access computer lab at the university) while they were carrying out their translation assignments for a medical translation module. After the completion of each piece of translation, in-class discussion would take place, focusing on potential translation difficulties encountered during the translation process. Discussions about web search occurred organically in this context although it was not a subject that was systematically taught in this module nor in any other modules. A quasi-naturalistic design was favoured in this study partly because it was aimed at capturing trainee translators' spontaneous web search process and partly because it minimised common issues of ecological validity often associated with translation process research, since subjects in this study were neither taken out of their normal place of work, nor were they asked to translate an unusually short piece of text out of context or in exaggerated font sizes.

At the beginning of the module, students were shown how to use the relevant software, i.e. BBFlashback Express or QuickTime (if using Mac) to screen-record their translation and web search process. Students were encouraged to verbalise their thoughts wherever possible, with the audio recording function being enabled in the software so that their verbalisation could be recorded alongside their on-screen behaviour. However, it is important to acknowledge that due to the quasi-naturalistic design, disparity exists in terms of the quantity of verbalisation, as some students did not verbalise consistently throughout their translation process and in some cases, verbalisation was minimal. It is for this reason that screen recording remains the primary data source for the present study. If verbalisation was taken into account, this would be pointed out specifically in the analysis. Screen recording provided two types of data for studying web search: one was screen-based activities (e.g. query typing, mouse-moving, and clicking behaviour, etc) and the other was timeline associated with them. The timeline was of course vitally important in depicting the attribute of efficiency in optimal web search.

Each of the three source texts belonged to a different medical genre. The first text was a medical case report, the second an academic article, and the third a clinical trial. Table 2 provides more details about these three texts.

6

INSERT TABLE 2 HERE

It is important to point out that this study focuses on web search processes associated with specific 'Rich Points' (PACTE, 2011, p. 41) in the STs rather than the translation process as a whole. In PACTE's study (ibid), Rich Points were referred to as specific ST segments that contain translation problems. For the purpose of the present study, I identified a number of ST terms as Rich Points in the three translation tasks. These Rich Points were identified largely based on post-task in-class discussion where locating TL equivalents on the web appeared to be particularly problematic for many trainees. I name the web search process associated with each of these Rich Points a 'web search episode'. Each web search episode commences with a query action associated with a particular Rich Point and ends with the searcher moving onto actions unrelated to the Rich Point (although occasionally, a searcher might move on and then decided to return to the same Rich Point. I consider this to be a second leg of the same web search episode). In total, 9 Rich Points were identified for analysis among the three STs (as seen in Table 3).

INSERT TABLE 3 HERE

In total, 97 web search episodes were transcribed. The transcription scheme largely follows the one used in previous studies, such as XXX and Enríquez-Raído (2014), although other metrics were added. Worth noting here is that almost two thirds of the web search episodes (60 out of 97 in precise) consisted of using one or two online dictionaries only. This was a consistent finding also reported in both XXX and Gough's (2015, p. 157) study, suggesting that the use of online dictionaries remains a particularly dominant source of information not only for trainee translators but also for professional translators. Gough (2015, p. 166) spoke in this case of a 'one-step research episode', meaning the web search process involving either a simple dictionary query, or one additional Google query in addition to the dictionary query. It has to be pointed out that the bulk of the analysis and discussion in the present study will be concerned with web search episodes that involve at least three or more steps, rather than 'one-step research episodes' (ibid).

4. Results and discussion

4.1 Unsuccessful web search episodes

This section will focus on what characterises unsuccessful web search episodes. For operational purposes, an unsuccessful web search episode is defined as a searcher not being able to locate a satisfactory TL term. As mentioned earlier, in most unsuccessful web search episodes, trainees tended to give up searching fairly quickly, posing one single ST term as a query either in an online dictionary or a search engine. This was probably because they either accepted whatever they found in the online dictionary as a TL equivalent without double-checking it. Or, if no entry was found in an online dictionary, they might decide that a better

7

option was to coin a new TL term. A typical example of this can be seen in Table 4. The Rich Point here was 'caput and moulding'. This term was associated with the swelling and shaping of a baby's scalp caused by the pressure against the birth canal during labour. There are several possible TL equivalents for this: 产瘤, 胎头水肿, or 先锋头. In Subject P10's case, none of these TL equivalents were found. S/he briefly posed parts of the ST term as queries in Baidu (the most popular Chinese search engine) and in Dictionary.com in quick succession. Then, s/he decided to coin a new term.

INSERT TABLE 4 HERE

Apart from the above example, lots of endeavour was also found in other unsuccessful episodes. Table 5 shows a typical example of this. This example was associated with the Rich Point, 'station 0'. This is an obstetrics term that proved to be very difficult to locate in the TL. It indicates a particular descending position that aligns with a mother's pelvis when a baby's head moves down in the birth canal during delivery. Its TL equivalent is 坐骨棘连线 齐平 or 坐骨棘平面齐平 (Literal translation: pelvic bone alignment).

INSERT TABLE 5 HERE

First of all, with regard to query duration, Subject P13 spent seven and a half minutes approximately, which, comparatively speaking, is a lengthy period of time to be spent on one Rich Point. S/he posed 13 queries in total using two types of web resources exclusively (i.e. Baidu and Zh.wikipedia). There were several interesting features worth pointing out here. First of all, Subject P13 repeatedly used the same web resources in this search episode. In fact, s/he only used a few habitual web resources throughout all her/his translation process. At the first glance of average query duration, Subject P13 spent an average of approximately 30 seconds glancing through search results before posing yet another new query. There was a high level of query abandonment, meaning no clicking of search results after new query submission, indicating a sense of dissatisfaction. In fact, a closer look at the timeline shows that sometimes Subject P13 spent a mere three to eight seconds glancing through search results (SERPs) before abandoning it. There seemed to be a great deal of urgency to produce more and more queries, and presumably the subject was aware of the length of time being spent on this single Rich Point. This was probably why very little time was actually spent on assessing or judging the effectiveness of search results, let alone clicking on them. In other words, the average trail length was minimal. As seen in Table 5, at the timeline of 15:38, Subject P13 clicked on the 1st search result in Baidu and then a mere 11 seconds later (metric: dwell time), s/he decided to move onto Zh.Wikipedia to pose yet another query.

In terms of the query formulation, the queries were either the ST term and/or a rough description of the ST term in the TL. It was clear that Subject P13 obtained a reasonable understanding of what 'station 0' meant probably based on the ST context, since all her/his queries were related to foetus' or baby's head position. Many near-identical queries were posed one after another and with very short intervals between them. In other words, there was

8

no substantial query re-formulation. This can be seen, for example, in the queries changing from 'station 0 婴儿头部位置' (literal translation: station 0 baby head position) at timeline 14:28 to 'station 0 胎儿头部位置' (literal translation: station 0 foetus head position) at 14:59 and then to 'station 0 胎儿位置' (literal translation: station 0 foetus position) at 15:02. As mentioned before, all the queries posed in Baidu seemed to be on the same line of baby/foetus (head) position throughout the whole web search episode. In the end, after such a lengthy and yet futile endeavour, Subject P13 decided to end the search episode and leave this term untranslated in the TT.

4.2 Successful and more optimal web search episodes

In contrast to the example above, Subject P1 manage to locate the appropriate TL term for 'station 0' relatively swiftly and painlessly. This more optimal web search episode is seen below in Table 6.

INSERT TABLE 6 HERE

Subject P1 started off by using an online dictionary, i.e. Bing Dictionary, to check for the term 'station 0', clearly without satisfactory results being found ; s/he quickly moved onto Google, using the same query. Then, interestingly, s/he decided to leave 'station 0' untranslated for the time being and continued to translate the ST. Up until this point, this appeared to be an unsuccessful 'one-step research episode' (Gough, 2015, p. 166).

However, when s/he came back to revise her/his translation the next day, s/he posed the query, '分娩 station 0' (literal translation: labour station 0). S/he then slowly move the mouse to hover over several snippets[2], indicating s/he was processing them, while verbalising '什么位置?' (literal translation: 'what position is this?'). After reading through these snippets, s/he decided to click on the 4th search result and read its contents at length (dwell time: 1:13) before exclaiming that she found the TL term.

Another example demonstrating a successful web search episode was that of Subject P2 (see Table 7). Like Subject P13, Subject P2 spent approximately seven minutes and posed 10 queries in total. In spite of a relatively long query duration (overall), at timeline 12:36, which was about two minutes after the first query, Subject P2 had already located the TL term. However, s/he was very cautious and decided to conduct further web search to confirm the TL term.

INSERT TABLE 7 HERE

---

[2] In IR research, "snippets" refer to captions and content summaries being displayed alongside individual search results (or webpages) in SERPs. (Gomez-Nieto et al., 2014, p. 457)

9

Interestingly, at the beginning, Subject P2 verbalised that s/he decided to pause the recording so as to concentrate on reading about the process of giving birth. S/he then came back to pose more queries for 'station 0', which were mostly near-identical, somewhat similar to those posed by Subject P13. But instead of rushing to pose more queries, Subject P2 actually took her/his time going through the contents of many search results, including the 1st, 2nd, 5th and 6th search results after queries.

In the cases mentioned above, in Table 7 and in many successful episodes, subjects tended to spend more time reading search results in SERPs, presumably making a more determined and conscious effort to understand and extract relevant information. Clicking and processing search results in SERPs beyond the top three constituted relatively unusual web search behaviour among all the subjects in this study. This is also true in IR literature in general, since 'position bias' (White, 2016, p. 65) is recognised as a common occurrence in web search behaviour, meaning, web searchers typically process their search results from the top-down position and, more importantly, they do not normally process beyond the first several search results from the top. Therefore, those search results being positioned at the top tend to catch searchers' eyes more. However, ironically, findings in the present study suggest that clicking and processing information beyond the first few results in SERPs seemed to be a winning formula to locate hard-to-reach TL terms. The actual number, length and (re-)formulation of queries (i.e. query-related metric or primary actions) became almost irrelevant. Obviously, this could be because most subjects did not use Boolean symbols or complex syntax in their queries. In fact, most queries were made up of either a ST term or a rough TL description of the ST term, or a combination of both. This is very much in line with the findings concerning query formulation in XXX. It is perhaps understandable that if the subjects' main aim was to locate TL terms, they would pose queries consisting of ST terms or possible TL terms. However, it was difficult to judge whether the use of more complex syntax in queries would have aided their searches since almost none of the subjects chose to do this. Again, as mentioned before, minimal use of complex syntax or Boolean symbols in queries is a stereotypical feature of web search in general (Jansen and Pooch, 2001, p. 240).

To reiterate, in contrast to the unsuccessful and ineffective episodes, a striking characteristic of successful or more optimal ones was that trainees took the time (longer dwell time) to read and assess SERPs. In other words, SERPs were processed in more depth rather than superficially. As search results for each query were assessed, trainees were then more likely to create a more tailored subsequent query, rather than posed more near-identical queries repeatedly. For example, after spending about 1 minute to read the 5th search result at timeline 12:36, Subject P2 posed a subsequent query, '胎头下降程度的表示方法' (literal translation: the way in which descending levels of foetus's head is shown), whereas all the previous queries had basically been 'station 0' in various forms. This suggests that follow-up queries did take into account of previous search results and queries. Instead of starting from scratch each time, subsequent queries built on previous queries and search results, which probably brought Subject P2 closer to locating a satisfactory TL term.

Taking a break also seemed to contribute to successful and more optimal web search episodes. This was true whether the student decided to move onto another term worth searching for, or simply decided to have a coffee break, or even left the whole web search episode until the next day, as is the case in translation revision process where 'drawer-time' XXX acts as a kind of buffer or mental refreshment, which appears to be a good strategy to manage potential frustration or mental fatigue experienced by translators.

4.3 What contributes to an unsuccessful web search episode

As mentioned in Section 4.1, many unsuccessful episodes were the result of 'one-step research episodes' (Gough, 2015, p. 166), and often searchers gave up searching fairly quickly or were not persistent in their search. However, there were also occasions where evident persistence still failed to yield satisfactory results. In many of these cases, despite many queries being posed, SERPs were not processed in detail. Little engagement with the actual SERPs was characterised by minimal dwell time and trail length, i.e. minimal secondary actions. There are several interesting aspects worth pointing out here. First of all, in most unsuccessful episodes, searchers hardly clicked and let alone investigated search results in SERPs. In other words, query abandonment was a vivid feature. There was also heavy reliance on snippets in SERPs. Such over-reliance suggests that the trainees might not be aware that snippets themselves were not all created equal and some of them were potentially misleading or unreliable. (Gomez-Nieto et al., 2014, p. 457). In addition, when search results were occasionally clicked and read, they were mostly limited to the first three in SERPs. Perhaps the trainees did not venture into search results further because they assumed that the most relevant search results would be listed at the top. There was a lack of awareness that commercial websites actually pay search engines to have their webpages ranked at the top of SERPs. It was no coincidence that many search results listed at the top of SERPs were commercial websites, whereas specialist medical websites may not be found near the top. This probably contributed to the reasons why specialist TL terms sometimes proved hard to locate. The combination of these factors meant that some trainees' web search endeavours were futile and often accompanied with frustration.

4.4 What strategic web search behaviour should trainees be engaged in

To summarise, one of the most striking difference between successful (or more optimal) and unsuccessful episodes was related to 'secondary actions', i.e. engagement with SERPs, rather than 'primary actions', i.e. query-related actions (White, 2016, pp. 21-37), whether this was by means of clicking beyond the top three search results or taking the time to process potentially relevant contents in selected search results. An added benefit of in-depth processing in SERPs meant that it enabled the trainees to formulate a more targeted query, based on existing search results. In-depth engagement with SERPs is very much reminiscent of 'the explorer', a type of web searchers that White (2016) and Gough (2015) both mentioned in their studies. The 'explorer' is not only a web searcher who engages with SERPs but also a web searcher who 'branches out more' or 'explores more' (so to speak) in their searches (Gough, 2015, pp. 229-237). In other words, they do not limit themselves to a linear or predictable path in web search. Strategically, this means that trainees should probably be more encouraged to explore unexpected places in SERPs since this might be helpful to locate a particularly difficult (to find) TL term on the web.

This lends itself to the notion of efficiency and effectiveness in the present study. Clearly, there is a dilemma here. The present study found that successful episodes clearly benefit from more in-depth and exploratory types of search behaviour. But this takes time and effort. Presumably, in both successful and unsuccessful episodes, the trainees' intension was the same: to locate TL terms as soon as possible. Yet, while some trainees chose to spend their time posing many simple and near-identical queries in quick succession, leaving relatively little time in assessing SERPs, others channelled their time and effort into examining SERPs. As it turned out, the latter approach paid off in terms of effectiveness, while the former became a false economy in efficiency. In the end, it was not necessarily the amount of time spent overall, but how and where it was spent that seemed to offer optimum return in web search. Finally, in cases where particular intense search episodes caused frustration, having a break (of whatever form) might be a good strategy to manage the web search process and offer searchers a second chance to re-visit a sticky point so to speak.

5   Conclusion

Unlike previous studies, this study did not focus on individual idiosyncrasies of web searchers, but on more optimal (or successful) vs unsuccessful web search episodes using an evidence-based approach. It aimed to find out why successful and unsuccessful episodes occur and how trainees can conduct their web search in a more optimal manner. The main finding was that exploring and processing SERPs at a deeper level can be an effective strategy to satisfy a rare information need, particularly in the context of harder-to-reach TL terms. There was little evidence, nevertheless, to suggest that extensive query re-formulation alone makes much impact to the success of web search. This finding appears to challenge previous studies (as seen in Enríquez-Raído 2011, 2014 and XXX) where query construction and formulation often dominate. This may suggest a slight shift of focus from primary actions to secondary actions in the field of web search investigation. However, it has to be pointed out that the finding of the present study is based on evidence from terminology search in a specialised field of translation (i.e. medical translation), which may not apply in web search for other information purposes in translation, or translation of non-specialised texts.

In terms of implications for information literacy, the findings of my study reconfirm that understanding about how search engine works is important (Pinto and Sales, 2007, 2008a, 2008b; Sales et al 2018; Enríquez-Raído 2014). In particular, better understanding of how SERPs are ranked and how snippets are produced will be of great help particularly in relation to how trainee translators conduct their secondary actions in web search. My findings also suggest that it is important for trainees to think about how they manage their time and effort in web search, including learning to manage affective factors (i.e. frustration and mental fatigue) in web search, since these affective factors may have negative impact on web search. It is hoped that findings such as these will assist researchers to further develop evidence-based curriculum and pedagogy for information literacy. Future studies could also focus on how and to what extent negative emotions may impact web search process and behaviours.

References:

Austermühl, F. (2014). *Electronic tools for translators*. New York: Routledge.

Bilenko, M., & White, R. (2008). Mining the search trails of surfing crowds: identifying relevant websites from user activity. In *Proceedings of the 17th International Conference on World Wide Web, April 21 - 25, 2008, Beijing, China* (pp. 51-60). New York: Association for Computing Machinery.

Enríquez-Raido, V. (2011). Developing Web Searching Skills in Translator Training. *Revista Electrónica de Didáctica de la Traducción e Interpretación*, *6*, 60-80. doi:10.24310/redit

Enríquez-Raido, V. (2014). *Translation and Web Searching*. New York: Routledge.

European Master's in Translation (EMT): Competence Framework. (2017). Retrieved from https://ec.europa.eu/info/sites/info/files/emt_competence _fwk_2017_en_web.pdf

Gomez-Nieto, E., Roman, F., Pagliosa, P., Casaca, W., Helou, E., Oliveira, M., & Nonato, L. (2014). Similarity preserving snippet-based visualization of web search results. In *IEEE Transactions on Visualization and Computer Graphics 20 (3)* (pp. 457-470). NY.

Göpferich, S. (2009). Towards a model of translation competence and its acquisition: the longitudinal study TransComp. In *Göpferich, Susanne, Jakobsen, Arnt Lykke and Mees, Inger M., eds. (2009): Behind the Mind. Methods, Models and Results in Translation Process Research. Copenhagen Studies in Language 37*. Copenhagen: Samfundslitteratur.

Gough, J. (2015). *The patterns of interaction between professional translators and online resources* (Unpublished doctoral dissertation). University of Surrey, Guildford, United Kingdom.

Jansen, B., & Pooch, U. (2001). A review of web searching studies and a framework for future research. *Journal of the American Society for Information Science and*

Technology, *52*(3), 235-246. doi:10.1002/1097-4571(2000)9999:9999<::aid-asi1607>3.0.co;2-f

Jansen, B., Spink, A., & Saracevic, T. (2000). Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing & Management*, *36*(2), 207-227. doi:10.1016/s0306-4573(99)00056-4

Kelly, D., & Belkin, N. J. (2004). Display time as implicit feedback: understanding task effects. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 377-384). Sheffield, U.K.

Khabsa, M., Crook, A., Awadallah, A. H., & Zitouni, I. (2016). Learning to account for good abandonment in search success metrics. In *CIKM'16: Proceedings of the 2016 ACM Conference on Information and Knowledge Management*. Indianapolis, IN.

Li, C. (2009). *非文学翻译 [Non-literary translation]*. Beijing: Foreign Languages Teaching and Research Press.

Massey, G., & Ehrensberger-Dow, M. (2011). Technical and instrumental competence in the translator's workplace: using process research to identify educational and ergonomic needs. *ILCEA*, *14*, 2-12.

Mayer-Schönberger, V., Cukier, K., & Hogan, J. (2013). *Big data: A revolution that will transform how we live, work, and think*. Boston: Houghton Mifflin Harcourt.

Nogueira, R., & Cho, K. (2017). Task-oriented query reformulation with reinforcing learning. Retrieved from http://aclweb.org/anthology/D17-1061

PACTE group. (2011). Results of the validation of the PACTE translation competence model: translation project and dynamic translation index. In O. B.

Sharon (Ed.), *Cognitive Explorations of Translation* (pp. 30-53). London & New York: Continuum International Pub.

Pinto, M., & Sales, D. (2007). A research case study for user-centred information literacy instruction: information behaviour of translation trainees. *Journal of Information Science*, *33*(5), 531-550. doi:10.1177/0165551506076404

Pinto, M., & Sales, D. (2008a). INFOLITRANS: a model for the development of information competence for translators. *Journal of Documentation*, *64*(3), 413-437. doi:10.1108/00220410810867614

Pinto, M., & Sales, D. (2008b). Towards user-centred information literacy instruction in translation. *The Interpreter and Translator Trainer*, *2*(1), 47-74. doi:10.1080/1750399x.2008.10798766

Sales, D., Pinto, M., & Fernández-Ramos, A. (2018). Undressing information behaviour in the field of translation: A case study with Translation trainees. *Journal of Librarianship and Information Science*, *50*(2), 186-198. doi:10.1177/0961000616666131

Shei, C. (2010). *翻译与网络资源: 理论,标准,实务[Translation and the web: theory, practice and norm]*. Taipei: Win Join Books Co.

XXX

XXX

Silverstein, C., Marais, H., Henzinger, M., & Moricz, M. (1999). Analysis of a very large web search engine query log. *ACM SIGIR Forum*, *33*(1), 6-12. doi:10.1145/331403.331405

Spink, A., Jansen, B., Wolfram, D., & Saracevic, T. (2002). From e-sex to e-commerce: Web search changes. *Computer*, *35*(3), 107-109. doi:10.1109/2.989940

White, M. D., Matteson, M., & Abels, E. G. (2008). Beyond dictionaries. *Journal of Documentation*, *64*(4), 576-601. doi:10.1108/00220410810884084

White, R. W. (2016). *Interactions with search systems*. Cambridge: CUP.

White, R. W., & Drucker, S. M. (2007). Investigating behavioral variability in web search. In *Proceedings of the 16th Annual World Wide Web Conference* (pp. 21-30).

Wu, W., & Kelly, D. (2014). Online search stopping behaviors: an investigation of query abandonment and task stopping. In *Connecting collections, cultures and communities: ASIS&T 77th annual meeting, Oct. 31*. Seattle, WA: American Society for Information Science and Technology.

Table 1 Primary and secondary actions

Metrics types	Explanations
Primary actions: Query-related metrics	
Query (re-)formulation	How queries are (re-)structured or formulated.
Query duration	The duration (or the length of time) a searcher spends on a query.
Query abandonment	A searcher not clicking on any of the search results.
Secondary actions: Clicking & browsing-related metrics	
Dwell time	How long a searcher spends in a webpage or website before moving on.
Trail (length)	A temporally sequence of webpages visited (in a search episode).

From White, 2016, pp. 21-37

17

Table 2 Details about the STs

STs	Genres	Titles	URLs
1st text	Case report	A Tear During Delivery (translating from paragraph 3 onwards only)	http://www.medicalprotection.org/uk/case-reports-january-2014/a-tear-during-delivery
2nd text	Academic article	Nutrient Intakes from Food of Lactating Women Do Not Meet Many Dietary Recommendations Important for Infant Development and Maternal Health (Translating Section 2.2 only)	http://dx.doi.org/10.4236/fns.2014.517177
3rd text	Clinical trial	Gradual Egg-tolerance Induction in Hen's Egg Allergic Children (Translating detailed description and eligibility partially)	https://clinicaltrials.gov/show/NCT02487420

Table 3 Rich Points in STs

STs	Identified Rich Points
1st ST	Station 0 caput and moulding Neville Barnes forceps Saxthorph-Pajot technique
2nd ST	Primary study baseline visit physical activity (PA) level food group
3rd ST	Open provocation test

Table 4: Subject P10's web search episode

Time frames	Queries	URLs	Notes
05.50	caput	Baidu	
	moulding	Baidu	
	moulding	Dictionary.com	
07:05	STOP		Starting to coin a new term

Table 5: Subject P13's web search episode

Time frames	Queries	URLs	Notes
13:40	Station 0	Zh.wikipedia	
14:00	Station 0 胎位 (LT: foetus position)	Zh.wikipedia	
14:20	Station 0 胎位 (LT: foetus position)	Baidu	
14:28	Station 0 婴儿头部位置 (LT: baby's head position)	Baidu	
14:59	Station 0 胎儿头部位置 (LT: foetus head position)	Baidu	
15:02	Station 0 胎儿位置 (LT: foetus position)	Baidu	
15:19	胎儿位置 (LT: foetus position)	Baidu	
15:38	产前胎儿位置 (LT: foetus position before birth)	Baidu	Click 1st result: Baidu Baike 胎位 Dwell time: 11 seconds
15:49			Open zh.wikipedia.org
15:56	胎位 (LT: foetus position)	Zh.wikipedia	Click 难产 in Wiki Dwell time: 40 seconds
16:36			Backtrack to Baidu Baike 胎位 (previously clicked webpage)
17:26	胎儿头部朝上 临盆 (LT: foetus head facing upwards right before birth)	Baidu	
17:55	胎儿头部位置 (LT: foetus head position)	Zh.wikopedia	No results found in Wiki.
18:05	胎儿头部朝上 (LT: foetus head facing upwards)	Zh.wikipedia	No results found in Wiki.
18:40	胎位 (LT: foetus position)	Baidu	Reading Baidu content about "occipital"
20:11	STOP		Start to self-translate "station 0" into TL.

Note: Literal translation is abbreviated as LT in this and all the following tables.

Table 6: Subject P1's web search episode

Time frames	Queries	URLs	Notes
09:16	Station 0	Bing Dictionary	
	Station 0	Google	
10:40	STOP		Leaving "station 0" untranslated in TT, but continue to translate the ST and search for other ST terms.
Having a break and resuming this episode the next day.			
08:57	分娩 station 0 (LT: delivery station 0)	Baidu	• TAP verbalisation: "什么位置？" (LT: what position is this?)
09:15			• Click 4th result: 每日医读 www.wtoutiao.com/p/15dwbsa.html • Read this result in details. • Think-aloud protocol: "啊！找到了！" (TL: Ah, I found it.) • Dwell time: 1 min 13 seconds

Table 7: Subject P2's web search episode

Time frame	Query	Browser/URL	Notes
	生孩子 (LT: giving birth)	Baidu	Think-aloud protocols: 要花一点时间看生孩子的过程。3-5 分钟 (LT: Need to spend some time reading the process of giving birth. 3-5 mins)
	Pause		
10:19	station 0	Youdao	
	station 0	Baidu	
	0 station	Youdao	
	分娩 0 位置(LT: labour 0 position)	Baidu	
	分娩零位置 (LT: labour 0 position)	Baidu	
	分娩零位置 (LT: labour 0 position)	Baidu	
12:20	分娩 0 station (LT: labour 0 station)	Baidu	
12:25	Read/verbalise snippets of 1st result		
12:36	Click 5th result; read www.wtoutiao.com/p/15dwbsa.html		Think-aloud protocol: "哦，这里" (LT: Oh, it's here.) **Note: TL term was found.**
13:35	胎头下降程度的表示方法	Baidu	
13:48	Click 1st result		Note: read through the 1st result
13:51	Click 2nd result		
14:10	Click 6th result		
14:21	Read the snippet of the 2nd last search result on the 1st SERP		Think-aloud protocol: "真有这种说法" (LT: This expression really exists)
14:42	Read the content of the 6th result		
15:03	胎头下降程度 S+0	Baidu	
15:17	Click 5th result		
15:24	胎头下降 S+0	Baidu	
16:05	Click the 2nd last search result; www.dxy.cn/topic/21977262		
17:33	Copy and paste a few previously-read webpages to refer back later		