University of
BRISTOL

Translation Productivity

Tools, Tracking, Ethics

June 2022

Authors

Valentina Ragni
School of Modern Languages
University of Bristol
Email: v.ragni@bristol.ac.uk

Elisa Alonso
Departamento de Filología y Traducción
Universidad Pablo de Olavide
Email: elialonso@upo.es

Lucas Nunes Vieira
School of Modern Languages
University of Bristol
Email: l.nunesvieira@bristol.ac.uk

For further analysis of data collected as part of this investigation, see:
L.N. Vieira, V. Ragni, and E. Alonso, 'Translator autonomy in the age of behavioural data', *Translation Cognition & Behavior*, 4.1 (2021), 124-146. doi: 10.1075/tcb.00052.nun.

Contents

Executive Summary

This document reports on a study that examined perceptions of productivity and activity tracking. Sixteen professional translators took part, nine freelance and seven in-house. Participants integrated productivity tools into their work and reported weekly to the research team for a period of four months (16 weeks) in the first half of 2020. Translators rated their own productivity and the usefulness of activity tracking data, and left open comments about the experience each week. To contextualise their answers, translators also self-reported some quantitative data, including project sizes and text types. They were asked to report on projects they worked on in SDL Trados Studio and memoQ. To calculate their productivity, they used the in-built tracking feature in memoQ and the Qualitivity plugin in Trados Studio. After the study's 16 weeks, translators took part in one-off focus group interviews. The focus of this report is on the project data (Part I) and questionnaires (Part II), but where relevant we draw on comments from the interviews to complement the discussion.

A host of variables – often not easily quantifiable – were reported to affect productivity. These involved many ancillary tasks that had to be carried out outside the CAT tool and which are therefore inherently not trackable by these tools' productivity features. Overall, it was felt that only considering measures like speed or number of edits disregards these tasks, which risks providing an incomplete or potentially skewed picture of translation. In the productivity self-assessments, freelancers were less satisfied than in-house translators about their productivity, with lower satisfaction especially after the coronavirus outbreak. A key difference in perception was registered between internal or self-tracking of productivity and external tracking. While the former was considered a useful practice, translators unsurprisingly had many reservations about the latter. It was felt that translation vendors could use productivity data against individual translators, for example to negotiate price reductions. Cloud-based work was identified as having the highest potential for ubiquitous tracking and potentially unfair practices, especially if client-translator relationships are not trust-based. Based on these findings, we make five recommendations:

1. More clarity and terminological precision in discourses around productivity and activity tracking, to explicitly identify the stakeholders and perspectives adopted.
2. In cases where translators do not control the collection of productivity data (e.g. cloud-based work), more robust consent processes and transparency about exactly what kind of data is collected and why.
3. Further investigations into the legal domain, for example to assess whether existing legal frameworks like the General Data Protection Regulation could be used or built-upon to safeguard translators against possible negative consequences of activity tracking.
4. The recognition that the negative consequences of tracking are likely to be perceived more acutely by self-employed translators, for whom retaining control over their work is important for job satisfaction.
5. A wider recognition of translators as productivity tool users. We identify a market gap with untapped potential for tools designed and built with translators in mind.

Introduction

As the demand for translation grows globally, larger volumes and tighter deadlines are putting increased pressure on translators, vendors and buyers alike. In this scenario, efficiency and productivity considerations play a key role in allowing for timely deliveries and increased capacity. Moreover, the technologization and commodification of translation have meant that quantitative methods of assessing translation products and processes have flourished and are now readily available. Activity tracking is one such method, used not only in the translation industry but in many other areas. The European Union has funded projects involving tracking in health, security, electricity, and mobility,[1] for example. Discussions about tracking therefore have a wider relevance in the context of big data, artificial intelligence, and surveillance, and have become prominent since the Covid-19 pandemic given the privacy implications of contact-tracing apps and online communication technologies.

In the context of translation, activity tracking consists of collecting a range of behavioural data from translators while they work on a text, for example by recording time and keyboard activity. Many translation companies already use some form of tracking to assess the productivity of in-house translators. Some companies now track freelance translators as well, whether that be to calculate hourly pay or to monitor progress.[2] While some form of tracking may be justified in certain instances, the practice is controversial[3] and largely unregulated. Some online translation platforms already implement tracking by default without any opt-out options.[4] However, tracking could also have potentially positive applications, for example as a means for translators to gain awareness of their own working processes, or to guarantee in-house translators' rights, i.e. that their break times and work calendar are fulfilled. Despite these contrasting perspectives, surprisingly little research has looked at how activity tracking is perceived and how productivity tools are used by professional translators in their daily work.

This report starts to address this shortage in three ways. First, by gathering perceptions and opinions from professionals during and after first-hand use of activity tacking tools. Second, by assessing the usefulness of activity tracking data based on professional translators' usual workflows, preferred CAT environments and real assignments. Third, by tabling broader discussions about productivity, and the benefits and risks of activity tracking.

The study was preceded by a workshop organised by the University of Bristol in December 2019 for members of the Western Regional Group (WRG) of the Institute of Translation and Interpreting. The workshop introduced some productivity tools and their capabilities, described some of the measures commonly used to capture productivity, and initiated a discussion about the concepts of productivity

---

[1] See https://trackandknowproject.eu/

[2] See e.g. M. Panić, 'Synergium use case: DQF as a Reliable Risk Management Tool', https://blog.taus.net/synergium-use-case-dqf-as-a-reliable-risk-management-tool [accessed 14 May 2021]

[3] L.N. Vieira and E. Alonso, *The Use of Machine Translation in Human Translation Workflows: Practices, Perceptions and Knowledge Exchange*, Report (Institute of Translation and Interpreting, 2018), (pp. 15-16). https://www.iti.org.uk/resource/the-use-of-machine-translation-in-human-translation-workflows.html [accessed 04 March 2021]

[4] See e.g. J. Moran, 'Extensive Cat Tool Logging – Big Brother or Language Technology Evaluation Panacea?' In *ITI Research Network E-Book 2018: The Human and the Machine*, (Institute of Translation and Interpreting, 2018), (p. 7). https://www.iti.org.uk/resource/iti-research-ebook-translation-interpreting-2018.html [accessed 04 March 2021]

and effort. The report is organised as follows. In the Methodology, we describe participant profiles, tracking tools and productivity variables, followed by the project and questionnaire data collection processes. The results are then reported separately according to the type of data (Part I: project data; Part II: questionnaire data). Finally, we conclude the report and present our recommendations.

Methodology

Freelance and in-house translators were asked to use productivity tools in their daily practice over sixteen weeks. They reported on the nature of their weekly work by completing individual online spreadsheets (project details and quantitative data, Part I). They also completed a weekly questionnaire to provide their opinions (qualitative data, Part II). A copy of both is presented in the Appendix. At the end of this weekly data collection phase, we interviewed the company director for the in-house cohort to gain a better understanding of the company's structures and usual working practices. All translators were then invited to discuss the experience in a series of one-off focus-group interviews that took place in June 2020. Although the focus of this report is the project and questionnaire data, we occasionally draw on interview comments to complement aspects of translators' weekly submissions. Participant (P) labels consist of a number (e.g. P1) plus a freelance (F) or in-house (IH) cohort indicator. The study was independently reviewed and approved by the Faculty of Arts Ethics Committee at the University of Bristol. Translators renewed their consent at each weekly submission and for the final interview.

Participant profiles

The main criteria to participate in the study were that translators (a) already used Trados Studio and/or memoQ and (b) had regular work in these CAT tools (at least some projects each week). Sixteen professional translators took part. Seven of them worked in-house for a translation company in Spain, and nine worked on a freelance basis. The freelancers were recruited through the WRG and the authors' professional networks. The in-house participants were recruited within the translation company that collaborated on this project.

Table 1 Participants' age and years of experience by cohort

Age		
Freelance	Min	29
	Max	48
	Average	38
In-house	Min	22
	Max	32
	Average	27
Years of experience		
Freelance	Min	2
	Max	22
	Average	9
In-house	Min	1
	Max	9
	Average	4

Recruitment was staggered, so participants who started later continued to submit data until they reached 16 weeks. For most participants, reporting started between late January and early February 2020 and finished at the end of May 2020. All in-house professionals lived in Spain. Freelancers were based in the UK, Germany, Italy, Spain and China. In-house translators always worked into Spanish. Freelancers worked with a variety of language pairs, usually into or out of English. Across both cohorts, languages covered were Chinese, English, French, German, Italian, Russian and Spanish. Several translators worked from more than one source language into their native language, and only one translator reported working both into and out of their native language. No one reported being bilingual from birth. As shown in Table 1, freelancers displayed a wider age range (min-max: 29-48) compared to in-house translators (min-max: 22-32). Freelancers were also more experienced on average (9 vs. 4 years).

Productivity tools and variables

Participants used two tools to calculate their own productivity: the Qualitivity plugin for SDL Trados Studio (henceforth 'Studio'), and the productivity feature in memoQ. The former is a separate application to be downloaded and installed in Studio. The latter is in-built in memoQ but inactive by default, so it needs to be activated manually in the program's options.

In this study, the researchers had no access to participants' desktops. The translators tracked themselves via these tools, having complete control over their settings and uses. No information was automatically shared either: all data was recorded locally by the translators, who then manually reported a portion of the data via their spreadsheets. The information provided consisted only of *how* translators worked (see below) rather than for whom or with what texts. Translators were reminded that they should not share any details that could identify either the clients or the texts in any way.

Productivity tools can record a wealth of information about the translation process, such as number of edits, number and duration of segment visits, translation time and speed, and modified words. While we invited all participants to make free use of the tools and explore their capabilities beyond the requirements of the study, we asked them to record three productivity-related variables each week, of which we focus on two: working times and translation speed.[5]

Project data

Every Friday, translators reported their weekly data per project onto their online spreadsheets (see Appendix). Weekly project data included the three productivity variables as well as more general project-related information about CAT tools used, language combination, genre and text type, number of files, type of work (e.g. translation, review, post-editing), project size and whether the project involved translation memories (TMs), term bases (TBs), and/or machine translation (MT).

Since the core goal of the study was to gather participants' perceptions of activity tracking and productivity, the self-reported project data was not intended to undergo a predictive or generalisable analysis. Rather, the data complemented the questionnaires by providing background
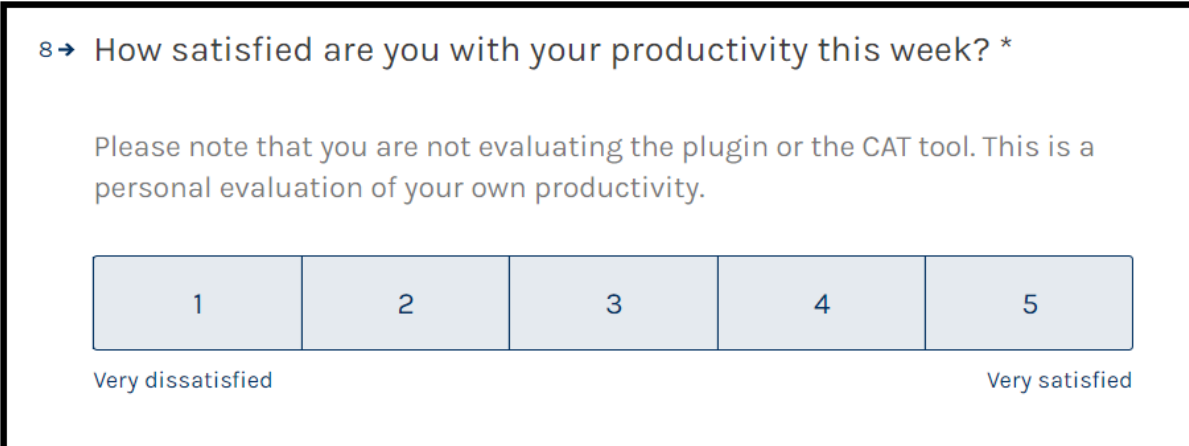
---

[5] Translators also reported the number of modified words, but this variable was calculated differently depending on the tool, so we do not report on it here to ensure cross-tool comparability.

information to contextualise translators' productivity self-assessments and open-text answers. If precise figures were impossible to obtain for any reason, translators were therefore asked to provide estimates. Moreover, if translators wanted to log projects they worked on outside memoQ or Studio, where productivity data was unavailable, estimating the information was also allowed.

The self-reported nature of the data meant the research team had to rely on translators to correctly extract the information from the tools. This process was inherently subject to events that on occasion prevented accurate reporting. There were missing data points caused by technology glitches, for example (see section 'Tool-related issues' below). Where translators reported figures that they questioned or did not deem trustworthy (for example because they were too high or too low), these were excluded from the counts during data analysis. Inevitably, self-reported productivity data of this kind only partially reflects the work undertaken by a translator and should therefore be approached with some caution.

Questionnaire data

The questionnaire was designed to gather translators' opinions, feelings and attitudes to productivity calculations and associated tracking tools. Each week, participants were asked to (1) indicate what independent uses they had made of the productivity tools, (2) assess their own productivity on a 1-5 scale and explain their rating (open question), (3) assess the usefulness of the activity tracking data on a 1-5 scale and explain their rating (open question), and (4) explain how they felt about tracking their own productivity (also an open question). Figures 1 and 2 show the 1-5 Likert-type scale and the corresponding open question used in the self-assessments.

8➜ **How satisfied are you with your productivity this week? ***

Please note that you are not evaluating the plugin or the CAT tool. This is a personal evaluation of your own productivity.

1	2	3	4	5

Very dissatisfied   Very satisfied

Figure 1 Example of a questionnaire item (rating in the productivity self-assessment)

Figure 2 Example of a questionnaire item (open question in the productivity self-assessment).

Part I – Project data

Project size

This section considers the total size, in source text (ST) word counts, of the projects translators worked on. Often, especially with extremely large projects, translators worked on a subset of the total ST word count. Because this is a *source text* word count, the figures therefore do not represent how much work the translators undertook but instead give an idea of the overall size of the projects across all task types (translation, proof-reading, post-editing, etc.). For the freelance cohort, the smallest project anyone worked on had 120 ST words and the largest 51,435. For the in-house cohort, the smallest project had 611 ST words and the largest 158,708. The numbers show wide overall variability in project sizes.

Figure 3 below shows the average project size per week, including projects that spanned over multiple weeks. Across nearly the entire duration of the study, project sizes were larger for in-house translators on average. The single exception was week 8, when freelancers worked on larger projects (15,153 average words) compared to their in-house counterparts (11,936 average words).
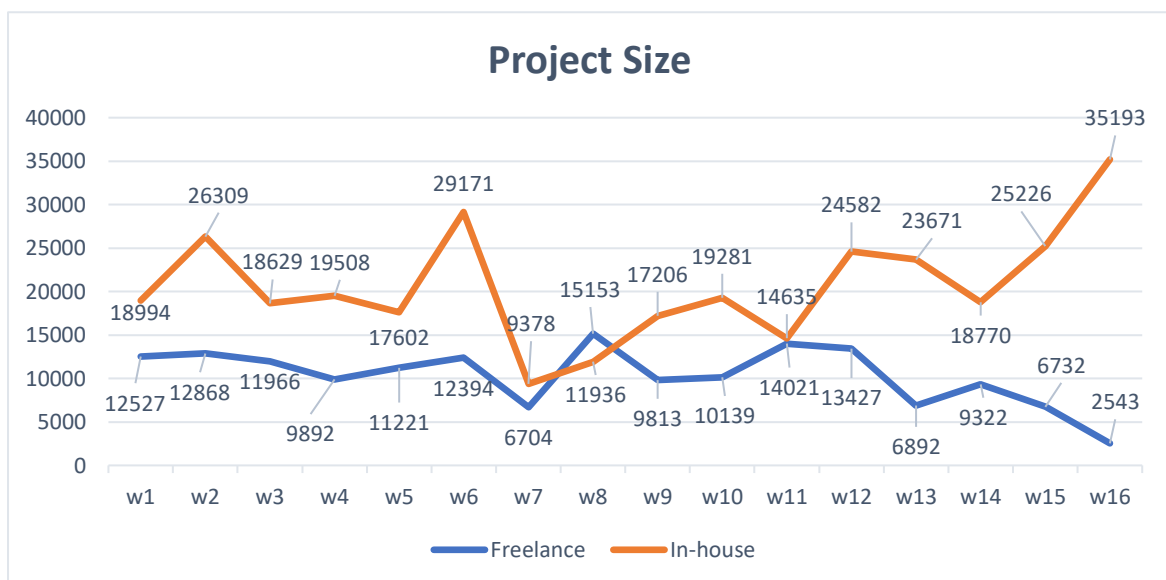

Figure 3 Weekly average project size in source text word count by cohort.

Reported working time

Table 2 summarises the reported total time spent on any tasks in a workflow (i.e. translation but also post-editing, proofreading, etc.). This is reported by cohort on a weekly basis. As previously mentioned, the study asked participants to log all Trados and memoQ projects. However, translators also worked on projects where activity tracking was not available, for example in cases where the project did not involve any CAT tools. Therefore, the values reported below are not necessarily representative of translators' *full* working week. On this basis, the figures show higher weekly variability in the freelance cohort, with a wider range of reported working times each week. Reported working week durations go from a minimum of just over one minute and nine seconds (00:01:09) to a maximum of just over 80 hours (80:00:25). Comparatively, in-house working week durations for the reported projects ranged from just over 30 minutes (00:30:15) to just under 50 hours (49:50:00). On average, the freelance and in-house cohorts spent comparable amounts of time working on reported projects throughout the 16 weeks of the study (11:00:13 and 11:27:34 respectively).

Table 2 Reported working time in seconds and hh:mm:ss (weekly values)

Reported working times		sec	hh:mm:ss
Freelance	Min	69	00:01:09
	Max	288,025	80:00:25
	Average	39,613	11:00:13
In-house	Min	1815	00:30:15
	Max	179,400	49:50:00
	Average	41,254	11:27:34

The numbers above show instances where freelancers worked extremely unsociable hours, including weekends. The freelance participant who worked the longest hours in a single week reported working 80 hours, which is roughly equivalent to 11 hours every day for 7 days.

Figure 4 shows the average reported working times per week. As can be seen, until week 12 reported tasks were longer for the freelancers on average compared to their in-house counterparts, a trend that is especially clear from week 5 onwards. However, this pattern changes in week 12 (marked by a dotted vertical line), when reported times become consistently longer for in-house translators until the end of the study.
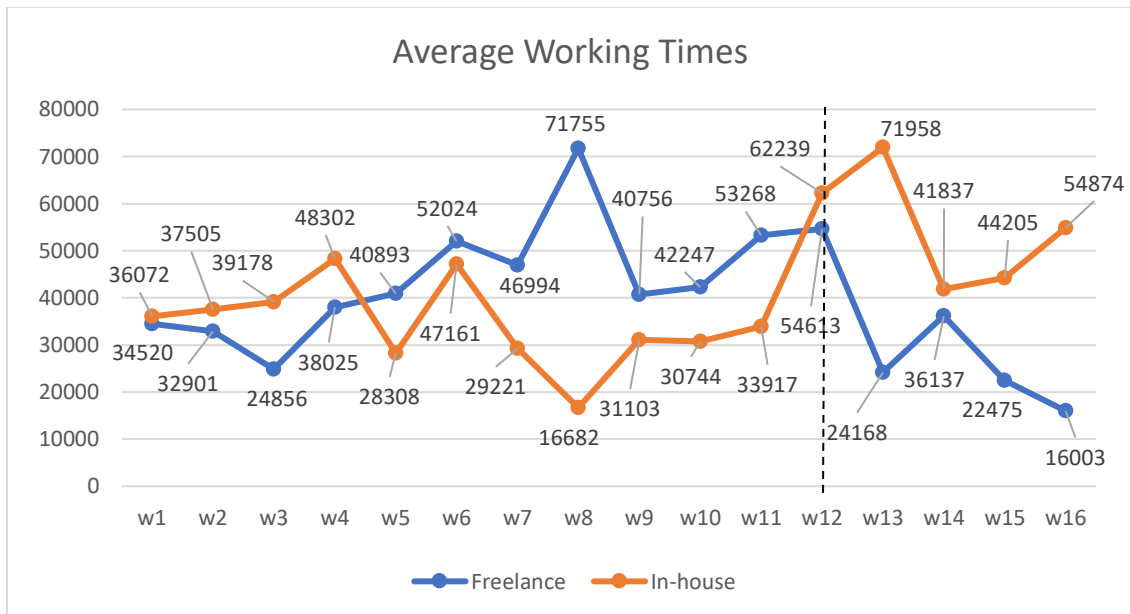
Figure 4 Average reported working time (in seconds) each week in both cohorts.

While the data does not allow us to prove causation, the change shown in Figure 4 is likely to reflect a differential effect of the Covid-19 pandemic on the two cohorts. Some freelance translators reported that they received little or no work in the weeks following the outbreak of Covid-19, which for most participants was in week 7 or 8. The graph shows that it took some time (2-3 weeks of regular work immediately after the outbreak) before work started to dry up for these freelancers. The open-text responses referring to Covid-19 in the questionnaire confirm this for some freelancers, but not for the in-house participants, whose job requests were not reported to diminish. In fact, in-house translators' reported times were longer on average after week 12 compared to previous weeks, with a peak of 71,958 seconds (approximately 20 hours) in week 13. This increase could be explained by an actual increase in workloads for the in-house professionals, but also by the in-house move to remote working, which was reported in the questionnaire to be somewhat difficult to adjust to at first, and therefore could have lengthened any tasks.

Translation speed

Table 3 displays weekly translation speed by cohort. By 'translation speed' we refer to speed of work across all task types carried out within the CAT tools (i.e. translation but also post-editing, proofreading, etc.). Translators reported speed data in Words Per Hour (WPH) for memoQ projects, and in Words Per Minute (WPM) for Studio projects. As most projects across both cohorts were carried out in Studio, all WPH figures were converted to WPM. Table 3 shows that, on average, translation speed was very similar between the two cohorts (54 vs 56 WPM).

Table 3 Translation speed in Words Per Minute (WPM) (weekly values).

Translation speed (WPM)		
Freelance	Min	19.66
	Max	383.955
	Average	55.89722
In-house	Min	14.25217
	Max	136.7888
	Average	54.34802

This overall similarity in translation speed is confirmed in Figure 5 below, where no clear trend can be subsumed across the 16 weeks of data collection. In exactly half the weeks freelance translators were faster on average (week 1, 2, 5-7, 9, 15-16), while in the other half the in-house translators were faster (week 3, 4, 8, 10-14). It therefore appears that, although the freelancers in this study may be dealing with smaller projects in terms of overall ST size, the translation speed of the two cohorts is comparable.

A note must be made at this point regarding the interpretation of the speed measures produced by the tools. Based on our tests in Qualitivity, when translators go back to an already-translated segment, make some further edits and confirm it again, their speed in this second visit may count as a separate data point in the WPM average, so WPM considers bursts of translation activity at specific points in time and does not necessarily reflect how long translators take to deliver a given document. By averaging different moments of the task, therefore, this measure does not correspond to a straightforward measure of total text volume per unit of working time and cannot be extrapolated to daily throughput. This means that a value of 54 WPM does not necessarily correspond to 25,920 words for an average working day of eight hours, which would be an extremely high figure even for tasks with machine translation and translation memory assistance carried out by experienced translators. The data suggested a similar interpretation for memoQ's WPH values, which on occasion were also extremely high (e.g. 6152 words per hour in one instance). Merging results from both tools at the minute level therefore did not seem problematic.

This way of calculating WPM introduced interpretability issues for the translators, who reported on multiple occasions not being used to this way of measuring their speed and preferring more straightforward measures of throughput. We revisit interpretability issues in Part II.
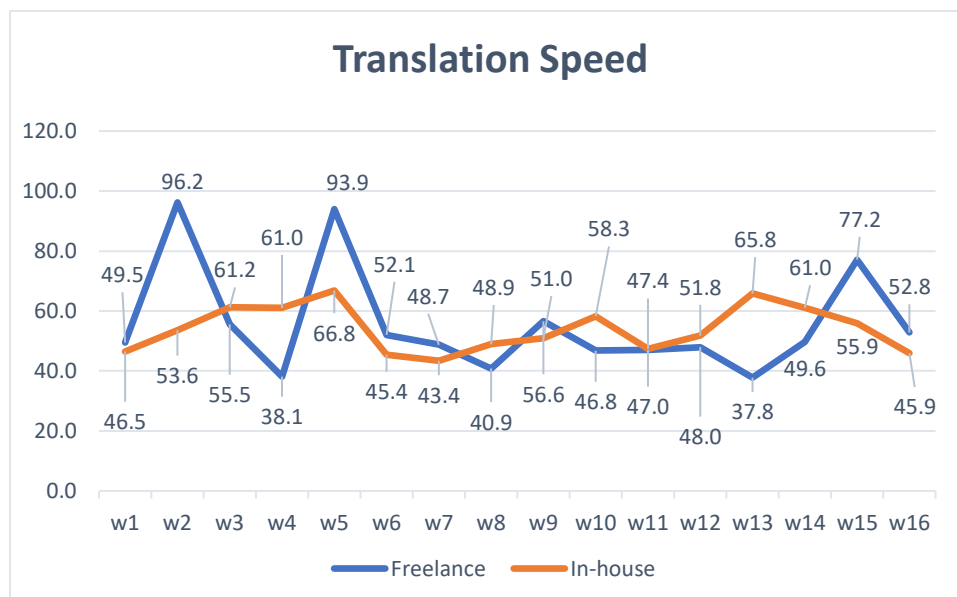

Figure 5 Average translation speed (words per minute) each week in both cohorts.

Untracked activity

Translators also reported a series of ancillary tasks that could not be analysed by the tracking tools but that nevertheless affected their productivity and could sometimes take a considerable amount

of time. In these cases, they reported the type of task and an estimate of the time taken every week. Two broad categories could be identified: (a) tasks that strictly speaking do not involve translation or other forms of text processing and (b) tasks that do.

Examples in the first category are the host of *administrative tasks* translators perform daily, usually per client or per project, such as liaising with project managers both via email and on the phone, uploading glossaries and/or past translations to client platforms, explaining translation choices, adding queries to client spreadsheets, signing and returning sworn statements, quoting, invoicing, and dealing with other forms of payment processing. P10-F also reported translating in a subtitling project that required numerous cross-checks of the accompanying video outside of the CAT tool (week 6), which she felt affected her translation speed. Other examples are *research tasks*, e.g. googling place names and photos for context, researching the company for which the translation job was carried out, researching products, studying client glossaries, guidelines and other materials, and seeking clarification of preferred translation approach with a vendor. Finally, *troubleshooting tasks* involving corrupted project packages that cannot be set up, glitches in clients' online cloud platforms, connectivity problems, and tracking issues.

Tasks in the second category required translators to perform some form of text processing outside of their CAT tools, for example dealing with poor quality .jpg documents where the OCR (optical character recognition) conversion does not work, which often involves manual formatting. Some translators also described a preference for proofreading texts outside the CAT tool, often on exported bilingual files, which is not accounted for by the tracking tools either.

Part II – Questionnaire data

Productivity tool uses

Every week, translators indicated how they used the productivity tools from the options presented in Figure 6. They were encouraged to explore the tools' capabilities independently beyond the scope of the study and could pick multiple answers if they used the features for different purposes on a given week.

11→ What productivity features did you use the most this week? *

Tick as many answers as applicable. You can also click on 'other' and type your own answer.

Choose as many as you like

A | creating a quote in Trados Studio / MemoQ

B | creating an invoice in Trados Studio / MemoQ

C | Measuring my translation speed by looking at measures such as words/minute, characters/second and the like

D | Examining my weekly client-level details, e.g. number of clients worked for, number of files/projects and word counts per client, etc.

E | Examining my weekly translation project-level details, e.g. number of files per project, word counts, type and number of TM matches, invoice amounts, etc.

F | Looking at document-level details, e.g. any by-segment information (time spent in each segment, segment revisits, number and type of edits, various edit distance statistics, etc.)

G | Other

Figure 6 Weekly questionnaire item on productivity tool uses.

Table 4 summarises how the participants engaged with the productivity tools. It shows how many times they exploited a given feature, expressed as a percentage of all responses. Translators used the last option ('Other', i.e. response 'G' in Figure 6) to mark weeks when they did not work with the tools, so these responses were excluded from the counts.

The first thing to note is that 'measuring translation speed' was chosen most often across cohorts (81% freelance and 39% in-house responses). Secondly, in-house translators exploited the project-level analyses more often compared to freelance professionals (33% of responses vs. 7%). Freelancers in this study carried out in-depth, within-document analyses relatively rarely (9% of all freelance responses vs. 22% in-house). On the other hand, they expressed a preference for comparing productivity variables across documents. For instance, they reported finding it useful to look at how their speed varied from project to project depending on the text type, project size, their area of specialisation and source language.

13

Tool use	Freelance	In-house
a. Quote	0%	0%
b. Invoice	0%	0%
c. Translation speed	**81%**	**39%**
d. Client-level details	3%	7%
e. Project-level details	7%	31%
f. Document-level details	9%	22%

Productivity self-assessment

Translators provided weekly self-assessments of their perceived productivity together with any corresponding comments. The productivity self-assessment was based on a 5-point rating scale (1 = very dissatisfied/unhappy; 5 = very satisfied/happy). Table 5 displays the distribution of the ratings between the two cohorts across the 16 weeks and shows that, overall, freelancers were less satisfied with their personal productivity compared with in-house translators. The most frequent satisfaction rating was 3 (neither dissatisfied nor satisfied) for freelancers and 4 (satisfied) for in-house translators. There was also a higher percentage of 1 and 2 ratings (very dissatisfied or dissatisfied) in the freelance cohort.

Table 5 Self-assessment of personal productivity (1-5 rating scale) across 16 weeks.

Rating	Freelance (n = 9)	In-house (n = 7)
5-rating	24% (n = 35)	29% (n = 32)
4-rating	27% (n = 39)	**47% (n = 53)**
3-rating	**34% (n = 49)**	14% (n = 16)
2-rating	8% (n = 11)	3.5% (n = 4)
1-rating	6% (n = 9)	2% (n = 2)
N/A	0.7% (n = 1)	4.5% (n = 5)
Totals	144 observations	112 observations

Somewhat worryingly, the results from Table 5 are in line with other reports that compare various aspects of translators' job satisfaction.[6] Translators reported fluctuations in workload and income as factors in their self-assessment of productivity. Lower satisfaction ratings were reported following the outbreak of Covid-19, in weeks when little or no work was undertaken.

Translators' comments suggested that speed and volumes of work were among the most important factors affecting productivity perception across both cohorts. Participants reported high satisfaction levels when they worked faster than expected and managed to process high volumes of words. They reported low satisfaction levels when they did not meet their personal speed and volume targets. However, speed and volume were not the only factors mentioned. Several translators also reported

---

[6] See e.g. J. Moorkens, 'Comparative Satisfaction among Freelance and Directly-Employed Irish-Language Translators', *Translation & Interpreting*, 12.1 (2020), 55-73 (pp. 62-65). doi: 10.12807/ti.112201.2020.a04 and H. Pielmeier and P. O'Mara, *The State of the Linguist Supply Chain*, (CSA Research, 2020) (pp. 62-65). https://insights.csa-research.com/reportaction/305013106/Toc [accessed 04 March 2021]

higher satisfaction levels when working on single, large projects compared to processing the same word count across several smaller projects, for example:

"I was able to work on a big assignment instead of many small projects. I definitively feel this helps productivity" (week 13, P05-IH)

This view is in contrast with current industry shifts towards the decomposition of multilingual projects and the 'slicing' of texts into smaller chunks to be split between several professionals on a piecemeal basis.[7] This preference for working with longer texts was often linked to the series of ancillary tasks discussed above. Since the administrative and research tasks associated with every project need to be carried out only once, longer texts leave the translator more time to concentrate on the linguistic task at hand, which in turn tends to correspond to higher satisfaction ratings in this study.

Factors perceived to negatively affect productivity include but are not limited to:

- low source-text quality;
- low MT quality in post-editing tasks;
- text type (e.g. marketing vs. technical);
- character limitations (e.g. in subtitle translations);
- distractions and interruptions to workflow;
- working from a second or third source language;
- technical problems, including connectivity issues when working with server-based TMs, computer crashes, shaky home network connections, and tracking issues;
- the hardware and software used, e.g. a laptop vs. a fixed PC, on cloud vs. desktop tools, in specific tools that they found clunky or new tools that they were not used to, for instance client-owned content management systems.

Importantly, life circumstances (e.g. bereavement, having a new-born baby, falling ill) and personal feelings were also often mentioned. For example, many translators reported low satisfaction levels due to 'lack of focus' (P4-F, P10-F), 'not being in the mood' (P7-F) and 'being frazzled' (P1-F), especially as the global coronavirus crisis unfolded. Across both cohorts, a difference between translators with and without family commitments also became evident during the pandemic, and productivity self-assessments were affected to a greater extent in those with caring responsibilities (e.g. for children or elderly relatives). School closures had a major impact in this sense. Some translators reported considerable difficulties in suddenly having to rearrange and split their time between translation work and home schooling:

"I wasn't very satisfied, because my work was interrupted multiple times (due to Covid-19 [which] induced home-schooling activities). Because of that, everything took longer." (week 8, P3-F)
"It's the 2nd week of working without any childcare due to covid-19 situation, so fitting in work and concentrating is more difficult than usual" (week 8, P2-F)

---

[7] J. Moorkens, '"A Tiny Cog in a Large Machine": Digital Taylorism in the Translation Industry', *Translation Spaces*, 9.1 (2020), 12-34.doi: https://doi.org/10.1075/ts.00019.moo

An intrinsic difference between in-house and freelance work was also highlighted by the pandemic. Unlike freelancers, who are used to working from home, in-house translators reported lower satisfaction ratings due to their move to remote working, e.g.:

"Due to circumstances, having to work from home, I feel my translation pace has slowed down a bit" (week 7, P01-IH).

Such was the strain associated with the sudden change in work setup that more than one in-house translator took a few days' annual leave to ease this adjustment period.

Aside of working on single large projects, factors perceived to positively affect productivity include:

- using one's usual CAT tool;
- learning to use a new tool faster than expected;
- working for particularly profitable clients;
- having a personal interest in the translation topic;
- dealing with high-quality source text and machine translation;
- working in one's area of specialisation;
- quoting on large projects; and
- meeting one's daily/weekly targets.

Notably, not working overtime and having a healthy work-life balance were also reported as having a positive influence on self-assessment rates. After the Covid-19 outbreak, just the fact of having work was reported as a determining factor in satisfaction levels regardless of job size or rates:

"Mainly relieved just to have work! (Covid-19)" (week 14, P1-F)

Activities like promoting one's business, updating one's website and making new clients also contributed to satisfactory self-assessments. Attending conferences and other continuing professional development events, on the other hand, can be perceived negatively as taking time away from translation during busy weeks. When work is scant, however, these events can also be perceived as productivity boosts. Finally, using the productivity tools for the purpose of this study was also mentioned as liable to influence self-assessments in both directions. Knowing that the time was ticking in the background in some cases helped translators to keep on task, maintain focus and avoid distractions, while in others it was perceived as a source of unnecessary additional pressure. While comments on the positive effects of tracking were reported especially at the beginning of the study, they tended to level off as time passed.

Usefulness of activity tracking data

Translators were invited to assess the usefulness of the information provided by the tools every week, using 1-5 rating scales (1 = useless; 5 = very useful) and open comments. Table 6 summarises their ratings across all weeks and shows a clear difference in opinions between the two cohorts. Freelancers rated the information as more useful. Their preferred rating was 3 (neither useless nor useful), against 1 (useless) for in-house translators.

Table 6 Usefulness of activity tracking data (1-5 rating scale) across 16 weeks.

Rating	Freelance (n = 9)	In-house (n = 7)
5-rating	15% (n = 22)	12.5% (n = 14)
4-rating	22.5% (n = 33)	26% (n = 29)
3-rating	**50% (n = 72)**	9% (n = 10)
2-rating	6% (n = 8)	17% (n = 19)
1-rating	6% (n = 8)	**31% (n = 35)**
N/A	0.5% (n = 1)	4.5% (n = 5)
Totals	144 observations	112 observations

We interpret this finding in light of the different levels of experience with activity tracking and productivity calculations between the two cohorts. In the focus group interviews, most freelancers confirmed that, before taking part in the study, they had very little experience with productivity tools. The few translators who used them primarily tracked the time they spent on different work-related activities with simple time trackers (e.g. freeware like Toggl or Clockify). In-house translators, on the other hand, had considerably more experience using an array of tools: different CAT-based tracking solutions, productivity variables calculated by their in-house translation management system, as well as external time trackers. Moreover, in-house professionals were more accustomed to discussing productivity with their line manager.

The comments that came with the ratings also help to contextualise these results. The information provided by the tools was deemed useful as a way of checking assumptions or beliefs, e.g.:

"Useful [as] a backup tool to provide data as support of theories (based on instinct)" (P1-F)
"I think that the productivity features are useful as hard data to back up a 'gut feeling' about how well you have done" (P3-F)

Translators also deemed the information useful as a source of insights into their own processes:

"Measuring time spent on the activities highlighted to me the way I work: I wasn't aware before that I spend so much time tweaking and fine-tuning the final text I deliver." (P3-F)
"I get best results in large projects than in smaller ones" (P06-IH)
"It gave me some insight in[to] the mistakes I make and why I make them" (P5-F)

Several freelance participants stated that the tools helped them to concentrate on the task at hand, e.g.:

"they helped me to gauge my speed and helped me stay on task, thus avoiding procrastination" (P7-F)
"It helps me stay on track and motivates me to work in short bursts without distractions" (P2-F)
"I think they can help me to focus on my work" (P4-F)

Translators also found it helpful to use the tools for forecasting purposes, i.e. to estimate how long a job will take, in order to be able to fit more work into one's schedule or organise one's working week:

"This week I've received +6k and I've been able to split it equally during the week according to a rough estimate of my translation speed." (P10-F)

Across both cohorts, there was a consensus that productivity tools work best with new clients or accounts, especially for quoting purposes in hourly paid jobs and to make initial operational decisions, for example to decide what tools to use with new clients:

"It helped me understand if Trados was the right CAT [tool] for a new client, since they also provide an online platform which, in the previous projects, proved more time consuming for me" (P5-F)

Translators often used productivity tools to analyse their own performance in relation to a specific factor, for instance comparing different weeks or different projects with the same text type:

"I found that my productivity is very similar for the same kinds of projects. Depending on the topic, I translate faster or slower" (P03-IH)

Tracking one's performance over longer periods was also deemed useful:

"I'm confident it will be more useful in the long run, i.e. quarterly or at the end of the year, when I will try to do a recap" (P10-F)

Some translators felt that activity tracking was especially useful in the context of revision projects, e.g.:

"Typically I have 2 activities in the report: one for translation, one for revision (the last revision I complete after letting the translation settle for a bit, before delivery). I'm used to only considering translation times, but having this distinction visually helps me to consider the amount spent in total on a project." (P10-F).

Productivity tools were also criticised for their complexity, however:

"It feels fiddly at the moment and it feels counter-intuitive to spend too long working it out considering it is a productivity tool" (P1-F)
"I would've found them [the results] more useful if I had had enough time to check them throughout the week." (P3-F)

Here we note that the study itself cannot be excluded as a potential factor in translators' perceptions of the time taken to access and process the activity tracking results, since they did this in some level of detail each week. In any case, it is also worth noting how translators' comments point to ways in which the tools may be improved and streamlined. Some of the tool-related issues encountered are described below.

Tool-related issues

Throughout the study, translators commented on several issues experienced with the tracking tools. The issues reported can be divided into three main categories: technical issues, granularity issues, and interpretability issues.

As for *technical issues*, among several problems reported by translators we note that during the study there was an issue with the pause feature of Qualitivity, which failed to stop the tracker during breaks:

"The information is useful but the Qualitivity complement is not discounting the "pause time" from the total so I have to do it manually" (week 10, P11-F)
"It doesn't work properly sometimes. On one of the projects I translated this week the timer wouldn't stop counting when I was inactive." (week 7, P03-IH)

The problem with the timer was subsequently corrected, but we note that this was among the reasons why part of the data had to be excluded from results reported in Part I. Specifically, whenever the translators reported these issues, we did not include the figures in the counts. If they reported the issue and provided an estimate, we used their estimate instead.

With regard to *granularity*, translators pointed out that they did not always deem the fine-grained level of detail produced by the tools necessary:

"I don't really get it why the plugin tracked that [many] details of my work – from the activity document, I can see it tracked the exact time, to the second, I worked on each segment, what was the segment before/after my work, etc. For me, a total number of words in a period of time will be good enough." (week 6, P6-F)

In relation to the Qualitivity plugin, the fact that the data was recorded for each translation session – or 'activity', as per the plugin's terminology – was also felt to make information about entire projects difficult to aggregate and obtain.

Regarding *interpretability*, translators, especially those in the in-house cohort, reported some degree of difficulty in trying to understand the wealth of information produced by the productivity reports:

"In my opinion, the features were really useful. However, this first week, I found it a bit confusing to detect the right details in the reports." (week 1, P08-IH)
"I don't know how to interpret the information well." (week 4, P01-IH)

It is worth noting that technical and interpretability issues at times seemed to conflate. As previously mentioned in relation to translation speed, translators referred to how words per minute or hour were too high. This on occasion may have resulted from technical problems, though as discussed above in the Translation Speed section was also the usual behaviour of these measures, e.g.:

"The thing about the WPH/WPM is that it doesn't really show a "real" number. For example, for one project on memoQ, I have a WPH of 2298.99, but I know that I def[i]nitely did not translate at that speed." (weeks 9-10, P01-IH).

Finally, many in-house translators felt that the productivity features were only relatively useful to them because they were already using other tools that satisfactorily calculated their productivity. Some participants stated that the tools they regularly used provided them with data that they found more interpretable because they were more familiar with the way the figures were produced.

Feelings about activity tracking

Translators were asked to describe how they felt about the act of tracking their activity and the use of productivity tools. A relevant insight provided by this section of the questionnaire is the fluctuation in personal feelings as the study progressed. Let us consider P02-IH (in-house) and P1-F (freelance) in Table 7 as examples.

Table 7 Some examples of participants' feelings towards activity tracking.

ID	Week	Comment
P02-IH	1	"I felt more in control of my workflow, since one sometimes feels that time flies by without being very productive (complex project and so on), and, with time and use, these tools will allow me to better control future deadlines."
P02-IH	2	"I'm always OK with using them because I like to measure my productivity, especially when I feel I'm lagging behind my usual productivity, so that I can analyse any loophole or waste of time."
P02-IH	8	"After having measured my productivity throughout the weeks, I feel that it is really useful to do that, but sometimes it can feel [like] a weight if you always think about it, or if you notice a decrease in productivity, as if there was a threshold you should always exceed."
P02-IH	9	"In general, I felt great this week because I was able to consistently increase my productivity."
P02-IH	16	"I have developed a love-hate relationship with this kind of tools; I don't always feel like using them, but in general, I like them because I like the feeling of working under pressure."
P1-F	1	"I found it focussed my mind on the task as I was conscious that if I took a 5 min break it would reflect on my figures. That felt good at points (i.e. stop being distracted and get on with it) but incredibly intrusive and unhealthy at others (i.e. I shouldn't second-guess the need for a comfort break!). I suspect that once I've got over my initial curiosity about my work speed and whether it matches my own belief of what it is, I might start to find this a bit too imposing. I'm still in the intrigued phase though!"
P1-F	3	"I'm beginning to be curious about how my measurements compare with other translators and whether I'd 'measure up' in the eyes of a client or how I would market myself depending on where I place (slow: methodical/careful? fast: efficient?). Which leads me of course to wonder how an agency would use that in my favour/to my detriment."
P1-F	9	"I knew I'd worked all hours (eves and weekend included), so it was interesting to see how they added up."
P1-F	10	"Useful, positive, first time I've really felt the detailed analysis could be relevant."

ID	Week	Comment
P1-F	12	"Still a little uneasy that a client would want to use it to measure (by which I mean judge) and cut costs. I still feel it is a good tool for me to use for my own purposes but not one I would want to cede control of. I find myself wanting to justify myself in this study by going 'yes I only did 3 hrs in Trados BUT...' etc. I did not become a freelancer/director of my own limited company to have to justify my time-keeping to anyone else!"

Throughout the study, P02-IH was overall very positive about the importance of calculating one's productivity and very rarely criticised the tracking tools. Nevertheless, when it came to describing feelings, they did express a contrasting sentiment towards tracking ("sometimes it can feel [like] a weight", "I have developed a love-hate relationship with this kind of tools"). P1-F's feelings also fluctuate throughout the study: they report being "intrigued" and "curious" when the experience is still a novelty, but as time goes by and they get used to self-tracking, ethical considerations about how tracking tools can be used by clients start to surface and tracking can feel "intrusive", "unhealthy" and "imposing". P1-F hints at how productivity tracking could lead to profiling and benchmarking ("how my measurements compare with other translators and whether I'd 'measure up' in the eyes of a client"), and explicitly mentions that activity tracking could be used for or against a translator.

P1-F also notes how tracking can be used as a form of judgment. Like P1-F, most translators in this study were aware of the potential issues brought about by tracking technology. Contrary to the tacit assumption most often implied when productivity is mentioned in the industry (more productive means faster, and faster is intrinsically good), P1-F points out how 'being slow' can be a positive characteristic if marketed as a proxy for attention to detail. P1-F also reiterates the importance of autonomy and control for professional freelancers, which might start to explain why tracking could be perceived as particularly problematic for self-employed translators. Like P1-F, many noted that a balance needs to be struck between quantification and the human nature of translation. All translators perceived their profession as a highly skilled, intellectual practice that requires specific cognitive skills and cultural knowledge. Many feared involuntary external monitoring and mandatory sharing could reduce translation to a measure of speed. Some freelancers underlined the contradiction of going freelance to be independent and avoid constant external monitoring, only to discover that this is what might be happening.

Using activity trackers internally for self-assessment during this study made participants aware of the level of detail these tools can collect. This, in turn, made many translators feel uneasy about the same tools being used externally. As P1-F aptly summarises: "I still feel it is a good tool for me to use, but not one that I would want to cede control of". The issue of control also came up during the interviews, where cloud-based work was identified as the place where control of tools and methods could most easily be taken away from translators, especially if client-translator relationships are not trust-based. In many personal experiences of cloud-based work prior to this study, freelance translators indicated that they were aware of being tracked but were not sure what exactly was tracked. One participant who worked as a reviewer in time-tracked projects stated: "I don't know how they use these data, to be honest. (…) I never receive any results or anything. (…) Reviewers don't really have much insight on what they [the company] are collecting or how they do that. But I

know that these data are closely monitored, because every once in a while, you get an email" (P5-F). In these emails, the participant explained, translators were asked to justify or elaborate on tasks where they had been faster or slower than average. These email queries sometimes also referred to the number of edits the translator had made, which suggests the data collected was more granular than just time spent on the task. The same participant mentioned a previous experience where a vendor they worked for decided to adopt a "tracker-based model" for calculating rates. The unilaterally imposed change was met with resistance from all translators, so the company had no choice but to negotiate. This example incidentally shows the power of collective bargaining, but it also reflects the intrinsically uneven distribution of power in the client-translator relationship, which should raise some concerns.

To summarise, participants expressed both positive and negative feelings about tracking during the study. Most negative feelings were associated with external tracking, and the loss of control over tools and methods that this might entail. On the other hand, internal tracking was associated with positive feelings of empowerment when translators learnt something about their working processes, or when they could use the tools fruitfully, for example to manage incoming deadlines or to aid the quoting process. It must be noted, however, that these positive feelings tended to be short-lived and to fluctuate as the study progressed.

Conclusions

Across cohorts, translators reported that several factors can affect their productivity, many of which are not trackable within a CAT tool. Although current translation industry practices tend to operationalise productivity as inherently trackable time or throughput measures, this study shows that, from the point of view of the translator, productivity is more than just time or speed. Such narrow conceptualisations, therefore – however convenient or practical – inevitably leave out important factors in a way that can disadvantage translators and should therefore be investigated further. On the other hand, as we have seen in the 'Productivity self-assessment' section (Part II), translators across both cohorts highlighted speed as one of the most important measures for them. Notably, therefore, although translators highlighted on many occasions the dangers of reducing translation to mere measures of speed or volumes translated, it is worth noting how they are still bound by a broader market logic that prioritises precisely these variables. When making the point that productivity is more than time or speed, translators thus criticise not only activity tracking but also this broader market mentality.

In terms of Project data (Part I), high variability in project sizes, working times, task types and requirements was recorded. Despite this variability, average translation speed was comparable across cohorts. Participants identified aspects of their daily work that were not tracked by the tools, such as research, administrative and troubleshooting tasks, and any instances that required them to work outside CAT tools (see 'Untracked activity').

In the Questionnaire data (Part II), translators elaborated on their experience. They discussed technical issues (e.g. tool glitches), pointed out that the measures are not always straightforward to understand, and explained that the information collected is more detailed than necessary. Collectively, this body of evidence suggests that companies are currently better served by

productivity tools than individual users or freelance translators. By examining translators' first-hand experience with some of these tools and their associated comments, this study helped to show how translators really work, which is hoped to provide a more holistic and realistic characterisation of translation processes. Another finding emerging from the questionnaire data is that, in this study, there was a perceivable difference in satisfaction ratings between the two cohorts. Freelance translators were on average less satisfied about their own productivity compared to the in-house counterparts (see 'Productivity self-assessment').

Overall, the study suggests that tracking translation behaviour is a practice with clear benefits and equally clear potential to be misused and to foster unfair practices. Some companies operating on the cloud are already tracking translators by default upon access to their platform (external tracking). In many cases, limited information about this behavioural data is shared with translators, if at all. Alongside issues of fairness, issues of ownership and consent therefore arise. In Europe, the 2018 GDPR regulations provide some protection for personal data and make it clear that the ultimate owners of this data are the people to which the data correspond. Future research is required on how far translators' productivity data may be covered by GDPR, however. Simply establishing ownership is not enough since issues of data collection, use, processing, management and storage also need to be considered. What are the exact purposes of tracking translation behaviour for a given vendor? What is the extent and the granularity of said tracking? Often, online platforms show a measure of speed and project completion percentage to the translators working on their cloud. Is there more data being collected behind the scenes? How is the data processed, used and stored by the vendor? Is it anonymised? If so, how robust is the anonymisation method? Is the data shared further with any third parties? If so, for what purposes? How do concepts like 'fair use' and 'legitimate interest' apply to tracking? Are privacy policies clear enough to answer these questions?

In recent years, translation industry stakeholders have started coming together with legal experts to address intellectual property rights and data protection laws in relation to practices like reusing translation memory content and crawling the web to collect data at scale.[8] Legal scholars have also started to examine translation industry practices and potential malpractices.[9] Institutions like the European Union have formally recognised the role played by online platforms in translation,[10] and some initiatives[11] may help private translation enterprises build data ethics into their business models. We argue that translation, law and ethics scholars should initiate joint discussions on activity tracking, capitalising on what has been learnt so far to provide answers to the many outstanding questions outlined above in a way that is positive for all stakeholders involved, including freelance translators.

Several translators in our study maintained that external tracking, if not transparent and consensual, can contribute to worsen existing imbalances of power in the client-translator relationship. Subject

---

[8] See e.g. *Language Data for AI, Report*, (TAUS: 2020). https://info.taus.net/taus-language-data-for-ai-report-download [accessed 04 March 2021]

[9] E.g. S. Yanisky-Ravid and C. Martens, 'From the Myth of Babel to Google Translate: Confronting Malicious Use of Artificial Intelligence – Copyright and Algorithmic Biases in Online Translation Systems', *Seattle University Law Review*, 43.1 (2019), 99-168. doi: http://dx.doi.org/10.2139/ssrn.3345716

[10] See https://ec.europa.eu/info/events/2020TEF_en.

[11] See e.g. https://www.ethicscanvas.org/.

to the legislation of different countries, being self-employed means that freelancers have already relinquished statutory rights enjoyed by their salaried counterparts such as paid holidays, sick pay and maternity or paternity leave. Many of them, including those in our study, did so because they enjoy the freedom that comes with their job. Some felt that the prospect of ubiquitous external tracking would eventually deprive them of that freedom. Therefore, if misused, activity tracking can influence how the translation profession is conceptualised in the long term, by shifting perception from an intellectually complex activity to a mechanical transfer of linguistic content that is easily quantifiable. In such a shift, each translator may well be considered 'a tiny cog in a large machine',[12] whose work is inherently replaceable to serve a productivity imperative. Contemplating these issues triggered feelings of uneasiness in some of the participants, who criticised practices like translator profiling and the unilateral imposition of tracker-based payment systems. Finally, issues of trust were mentioned when discussing personal feelings about tracking. Several participants highlighted that, when working for trusted clients who in turn trust them as translators, external tracking methods are not seen as necessary or useful and therefore are rarely applied.

Recommendations

Five key recommendations are put forward based on the data analysed and the conclusions presented in this report:

1. This study showed that a profound difference exists between self-tracking and external tracking from the point of view of professional translators, and that current industry definitions and operationalisations of these concepts tend to be biased towards private enterprises. Therefore, we call for more clarity and terminological precision in discourses around productivity and activity tracking. It is important to explicitly identify the stakeholder perspectives adopted, what is meant by productivity, and whose productivity is being considered.

2. Cloud-based platforms were identified by translators as the loci that are most prone to potentially ubiquitous, non-transparent, non-consensual external tracking. Until clearer regulations are introduced, cloud-based tools can indeed exacerbate the negative impact of activity tracking. We therefore call for more robust consent mechanisms when working online. Platforms should be more transparent about what kind of data is collected and why, and stakeholders should work together towards implementing ethical practices.

3. Crucially, since this study presented some evidence that tracking can foster practices that are unfair to translators and since activity tracking is so far *de facto* unregulated, we call for further investigations into the legal implications of this subject. For example, could existing legal frameworks like GDPR be used or built-upon to safeguard translators against possible negative consequences of activity tracking? As we have shown in the Conclusion, there are many grey areas that would benefit from collaborations involving legal experts and professional associations.

---

[12] Moorkens, *Translation Spaces*, 2020.

4. As shown by the questionnaire data, tracking practices are likely to affect freelancers and in-house professionals differently. We call for a recognition that the negative consequences of activity tracking are likely to be perceived more acutely by self-employed translators, who might have specifically chosen to 'go freelance' to avoid the kind of monitoring and control that activity tracking typically entails.

5. Finally, translators' assessment of the information provided in activity-tracking reports highlight several obstacles that undermine uses that translators themselves could make of the data. Translators raised issues with excessive granularity, interpretability and technical complexity. We therefore call for a recognition of translators as productivity tool users. Translators unanimously recognised the importance of productivity calculations and the tools' potential, so long as features, design and data structures were made more relevant to them.

Appendix

Spreadsheets

Figure 7 below presents the spreadsheet translators used to submit their weekly project data. Each translator had their own unique copy of this template, which they accessed online each week to populate the relevant tab with their project data. They were asked to create a name for each project (rather than use the projects' real names) and add the information at a project level with one project per row. Productivity variables are highlighted in red.

Figure 7 Project data spreadsheet template.

Questionnaires

All questionnaires were administered online through the Typeform platform. Below we report the questions analysed in Part II of this report.

Q8. **How satisfied are you with your productivity this week?** [1-5 scale]
Please note that you are not evaluating the plugin or the CAT tool. This is a personal evaluation of your own productivity.
Rating Scale: 1: Very dissatisfied; 2: Dissatisfied; 3: Neither satisfied or dissatisfied; 4: Satisfied; 5: Very satisfied.

Q9. **In relation to the previous question, could you explain why you are satisfied about your weekly productivity to the level you indicated above?** [open question]

Q10. **What productivity features did you use the most this week?** [multiple choice question]
Tick as many answers as applicable. You can also click on 'other' and type your own answer.
a. Creating a quote in Trados Studio / MemoQ
b. Creating an invoice in Trados Studio / MemoQ
c. Measuring my translation speed by looking at measures such as words/minute, characters/second and the like
d. Examining my weekly client-level details, e.g. number of clients worked for, number of files/projects and word counts per client, etc.

e. Examining my weekly translation project-level details, e.g. number of files per project, word counts, type and number of TM matches, invoice amounts, etc.

f. Looking at document-level details, e.g. any by-segment information (time spent in each segment, segment revisits, number and type of edits, various edit distance statistics, etc.)

g. Other

Q11. **In relation to the previous question, could you explain how you used these features of the plugin?** [open question]

What information were you trying to extract? For what purpose? Did you notice anything interesting or relevant about the tools that you would like to tell us about?

Q12. **How useful did you think the information provided by the productivity plugin was this week?** [1-5 scale]

Please note that you are now evaluating the tools you used, rather than your own productivity.

Rating Scale: 1: Completely useless; 2: Not very useful; 3: Neither useful or useless; 4: Useful; 5: Very useful.

Q13. **In relation to the previous question, could you explain why the productivity features you used this week in Qualitivity/MemoQ were useful/useless?** [open question]

What exactly did you like/dislike about these features and why?

Q14. **This week, how did you feel about using productivity tools and about productivity tracking more generally?** [open question]

This question is not compulsory, but we would be grateful if you could also share any feelings and/or emotions about this topic.