
Article:
Walker, C orcid.org/0000-0001-7949-5621 (2021) Investigating How We Read Translations: A Call to Action for Experimental Studies of Translation Reception. Cognitive Linguistic Studies, 8 (2). pp. 482-512. ISSN 2213-8722

https://doi.org/10.1075/cogls.00087.wal

Investigating How We Read Translations:

A Call to Action for Experimental Studies of Translation Reception

Dr Callum Walker
University of Leeds
c.m.walker@leeds.ac.uk

Profile

Callum Walker is a Lecturer in Translation Technology at the University of Leeds and Honorary Research Fellow at University College London, and has previously held lecturing positions at Durham University, Goldsmiths College University of London, and University College London. His research focuses on how experimental methods can be employed to investigate and measure stylistic equivalence between source texts and target texts, with a particular focus on the cognitive effort evoked in literary texts by stylistic language varieties. He is the co-editor of the Benjamins Translation Library volume *Eye Tracking and Multidisciplinary Studies on Translation* (2018, with Federico M. Federici) and has recently published a monograph entitled *An Empirical Exploration of Equivalent Effect in Translation: The Reader Experience of Literary Style* (Palgrave).

Investigating How We Read Translations:

A Call to Action for Experimental Studies of Translation Reception

Abstract

Since its inception, Translation Studies has hinged on theoretical concepts of effects and reception, with various reader-oriented notions such as equivalent effect, *skopos*, acceptability and adequacy, and user-centredness, to name but a few, having pervaded the discipline for decades. Despite this preoccupation with the phenomenology of translations, we still know very little about how translations are actually experienced – *written* translations especially. This article calls for an expansion of research into the reception and experience of source texts and their translations, reviewing the opportunities afforded by recent technological developments in eye-tracking, galvanic skin response sensors, echocardiogram monitors, and other multi-sensory devices. Using a short case study, a number of research questions and an outline of an experimental method are proposed to contrast the reading experience of two translations of the same source text, serving as a prompt for future research of this kind. By drawing inspiration from the few existing examples of research in this incipient paradigm and the considerations offered in the example, this article aims to stimulate future research to explore the vast untapped potential in this area and to arrive at a better understanding of the effects that different translation approaches yield and the potential variation in effects between source and target text.

Keywords: cognitive translation studies, reception, experimental study of translation, effects, reader response

Introduction

Considering the centrality of notions such as reception, perception, effect, purpose, and others to the discipline of translation studies (TS) since its inception, it is surprising that so little attention has been paid to the final part of the communicative transaction implicit in translation. If we do not consider *how* translations, in their various forms, are read, seen or heard, then how can we even judge the success of this enterprise? The very acts of reception, and the effects of translation , are of critical importance to the people that rely on translation, and yet, we still know relatively little about the cognitive, emotional, experiential, phenomenological, and other effects of translations – in all their forms – on readers, listeners and viewers.

Since the 'success' of a translation is reliant not only on the manner in which it is produced (the translation process), but also on the manner in which it is received and experienced (translation reception, broadly defined), experimental research dedicated to the production of translations and the reception of Audiovisual Translations (AVT) needs to be further extended and adapted to the reception of written translation. The aim of this paper is therefore to consider what experimental research can offer to traditional

theoretical concepts at the heart of TS, and to offer a means to examine the effects of translation in different contexts. Beginning with a conceptual discussion of the notions of 'reception' and 'experience', it sets out to explore some of the tools and instruments that could feasibly be used in experimental settings to observe or measure the moment-to-moment effects of translated texts. It closes with an example of possible research questions and a possible experimental design to explore different types of effects that may result from the choice of source text (ST) and target text(s) (TT). The article concludes with some observations on the importance of mixed-methods approaches and the role of TS in phenomenological research more generally.

The Lacuna

Over the last decade, TS has witnessed rapid growth in the number of experimental or quasi-experimental studies focusing on the translation process (see Olalla-Soler et al. 2020; Xiao & Muñoz Martín 2020: 2-4), a paradigm which is commonly referred to as translation process research (TPR). Worryingly, from the perspective of other experimental paradigms of TS, TPR has come to be almost synonymous with the broader umbrella term of cognitive translation studies (CTS) in spite of there being far more cognitive dimensions to TS than the translation process alone. TPR views translation as a cognitive-human behaviour involving planning, reflection, decision-making, and emotional responses (see Jakobsen 2017 and Jakobsen & Alves 2020). Using the technological advances afforded by techniques such as keystroke logging and eye tracking, among others, the TPR paradigm has yielded experimental methods and findings on matters such as the process of producing translations in different forms and directions, translation revision, and related tasks such as post-editing, source text and/or target text reading processes, interaction with the translator workstation, broadly defined, the characterisation of translator expertise from experimental data – a shift referred to as the 'predictive turn' in TS (Schaeffer et al. 2019) –, and methodological considerations more generally (Balling & Hvelplund 2015; O'Brien 2009, 2013; Teixeira & O'Brien 2018). This summary is but a snapshot of the current state of play in TPR. But what becomes apparent on reviewing this ever-growing list of TPR publications is the glaring absence that lies in its shadow. As the number of TPR publications continues to grow, the comparative absence of publications focusing on how the translated product is actually *received* by the reader or audience is brought into stark contrast.

Over the last decade, the field of Audiovisual Translation (AVT) has paid increasing attention to how AVT products are received and experienced, and, in fact, has a much longer tradition of empirical research with eye-tracking techniques in particular, with some of the earliest work dating back to the late 1980s (for example, d'Ydewalle et al. 1987). A lot of work in AVT reception (for overviews, see J.-L. Kruger 2018; Orrego-Carmona 2018) has focused on how readers process subtitles, with further attention paid to one- and two-line subtitles, the distribution of attention between subtitles and the audiovisual stimulus, and the cognitive effort involved in reading subtitles. Some studies have also triangulated eye-tracking fixation data with pupillometry, and electroencephalography (EEG), alongside self-reporting measures (for instance, J.-L. Kruger et al. 2013). The AVT paradigm therefore seems to be making far more positive strides to

complement our understanding of the translation process with an appreciation of how the product itself is received.

Yet, while AVT may be one of the common forms of translation with which large numbers of people interact on a daily basis due to the rise in video-on-demand platforms, this heightened attention in TS to the reception of AVT products is again, like TPR research, in stark contrast to attention paid to the reception of written translations, the very driving force of international business, diplomacy, policy-making, healthcare, research and development, and cultural exchange. It is this area – *written* translation – on which this paper focuses. The breadth and diversity of text types, purposes, functions, formats, and styles inherent in the written translation sphere represent a huge opportunity for enhanced research on the reception of translations, taking into account the specificities of different texts and forums of reception. This paper will engage specifically with the reading experience of literary texts in particular, but the underlying intention is to lay the conceptual framework, by using this one specific domain, to underpin experimental approaches to investigate the reception of different text types in different contextual settings.

In the *Handbook of Translation and Cognition* (2017), reception is not explicitly addressed in Alves and Hurtado Albir's (2017) or Muñoz Martín's (2017) concluding chapters on 'Moving Forward' in this field. Indeed, the only detailed treatment of textual reception in CTS in this *Handbook* comes in Haidee Kruger and Jan-Louis Kruger's chapter (2017) on 'Translation and Reception', in which they conclude that there has been 'limited empirical attention [to] the "effects" of translation in terms of the cognitive processing involved in readers' reception of translations' (Ibid.: 71). One of the few other comments on reception in the *Handbook* is a fleeting – albeit extremely noteworthy – remark by Hvelplund that 'research into the reception of translated text using eye tracking […] is largely uncharted territory, and a large research potential remains to be explored' (2017: 259). Reception only features in passing in various chapters of the recent *Routledge Handbook of Translation Cognition* too (Alves & Jakobsen 2021). Evidently, the relative absence of experimental research into the reception of written translation in such volumes highlights the pressing need for further research of this kind. and such research should be viewed as complementary to many of the forward-looking perspectives presented by Kotze (2020).

The Centrality of Reception to Translation

When we reflect on the theories that have come to the fore in TS over the years, one might rightly wonder why the reception of translations has not featured more prominently in empirical research over the last decade. After all, translation theories have hinged on notions of reception and effects since scholars began to theorise translation (see Brems & Ramos Pinto 2013; this argument is also made very convincingly in H. Kruger & Kruger 2017). Equivalent effect, for instance – popularised by Nida (1964) –, was an essential aspect of so-called 'dynamic equivalence', where 'the relationship between receptor and message should be substantially the same as that which existed between the original receptors and the message' (Nida 1964: 159). Toury's concepts of 'adequate' and 'acceptable' translations (Toury 1995/2012: 56-61) also suggest an implicit reception element, based around whether source or target culture norms prevail. Venuti's (2008)

advocacy of 'resistant' or 'foreignising' translation strategies to combat the traditional 'domesticating' approach also points to elements of reader-oriented norms and acceptance. Even in the functionalist paradigm, *skopostheorie* (Reiss & Vermeer 1984) and related approaches (Holz-Mänttäri 1984; Nord 2005) are all based on the functionality or adequacy of the translation from the perspective of the user, and how well the translation is deemed to satisfy this function in light of a particular goal. These same principles have been taken up more recently by Suojanen et al., who 'emphasize the central role of the user, or reader' (2015: 1). There has also been, beyond the aforementioned broader theories of translation, no shortage of discussion on different types of texts and the ways in which they engage readers for different purposes (Nord 2005; Snell-Hornby 1995).

Chesterman has arguably engaged with the matter of effects more than most, in his paper discussing 'translations as phenomena that have both causes and effects' (1998: 201). The causal dimension has been covered extensively in translation theory, but his claim that 'the decisions that translators make, and hence the translations they produce, have effects on the people that read them' (Ibid.: 219) has seen little empirical exploration. Chesterman describes an 'effect' as a 'change in mental state (emotional, cognitive, etc.) in the reader' – a 'proximate effect', to use his terminology, which he does not consider to be directly observable. 'Secondary effects' (or 'behavioural effects') are then defined as subsequent actions which can in some cases be observed, but in other cases not (e.g. 'an increase in knowledge' or 'aesthetic experience'). 'Tertiary effects' are considered to impact the target culture as a whole (e.g. intercultural relations, target language norms, etc.). Chesterman delineates three laws:

> [1] Law of heterogeneous effect: translations tend to have different effects on different people.
>
> [2] Law of changing effect: even with respect to a single reader, the effects of a translation change over time.
>
> [3] Law of multiple effect: even with respect to a single reader at a given time, translations tend to have more than one effect. (Ibid.: 220-221)

These laws touch on the crux of the argument made in this paper, namely that comparisons can be drawn between the effects on source text readers and the effects on target text readers, in terms of their similarity (whether intentional or otherwise).

While Chesterman's first law would seem to suggest that experimental studies will yield extremely heterogeneous results due to the wide range of different effects that different participants experience, I contend that variation in effects does not preclude experimental investigation or even wider generalisation (see discussion below). The second and third laws have further implications for the experimental design in terms of how to investigate more than a single effect, and how those effects might change at different points in time (pre-experiment, mid-experiment, post-experiment). The questions indirectly raised by these laws are addressed in more detail below where I discuss the proposed conceptual distinction between 'reception' and 'experience'. However, the caveat I would add to Chesterman's laws is that they still fall foul of the subjectivity problem; they are, after all, assumptions and – while logical – need to be testified on a wider empirical or experimental level before we can be more assured of their veracity. This call is precisely what

Chesterman addresses in the final pages of his paper, where he argues that researchers should test hypotheses about translation effects:

> I would make a plea for much more application of reader response theory, for more psychological studies of comprehension and readability with respect to various kinds of translations […]. (Ibid.: 227)

More than twenty years after the publication of this paper by Chesterman, it seems that very little has been done in this regard compared to other fields of translation studies.

Existing Experimental Research on Reception

While not wishing to diminish the value or scope of work conducted in the empirical domain of translation reception research, using methods such as surveys, questionnaires, think-aloud protocols, and similar approaches, the focus of this article is on *experimental* (or *quasi-experimental*) research methods to explore the reception of translations (and in some cases the corresponding source texts too).

There have been experimental studies of textual reading in TS, but such reading has usually been a precursor to a subsequent, typically process-oriented task, such as translation itself. For example, three chapters in Sharon O'Brien's edited volume (2011) address reading modalities in/for translation (Alves et al. 2011), the cognitive effort involved in reading, and subsequently translating, metaphors (Sjørop 2011; see also Sjørup 2008 and 2013), and how attention is distributed between ST and TT during translation (Jensen 2011). In a more widely known study, Jakobsen and Jensen (2008) conducted an eye-tracking study on how professional and student translators read for comprehension alone, how they read with a view to translation, how they read with a view to providing a spoken 'sight' translation, and how they read with a view to actually preparing a written translation. The findings revealed clear and progressive increases in the distribution of visual attention from one task to the next. Dragsted and Hansen's chapter in the same volume (2008) also investigates the coordination of reading and writing process in translation. All of these studies share a focus on reading specifically for the purpose of translation; they do not address the experience of reading *per se*, but rather how the process of reading interconnects with the process of translation production. As will become apparent below, experimental studies on 'reading for reading's sake' (e.g. reading literature) and even reading with alternative functionalist purposes (e.g. reading educational materials, reading advertising materials, reading a technical manual for subsequent use of the corresponding item) are few and far between.

One of the first studies investigating the reception of written translations using (quasi-)experimental methods is Haidee Kruger's eye-tracking research on the processing of foreignised elements in translated South African picturebooks (H. Kruger 2012; 2013). In this work, she focused on how specific translation strategies influenced the processing of and responses to linguistically and culturally foreign text, against the backdrop of linguistic hybridity in South Africa. Specifically, she was interested in the cognitive effort involved in the reading process, and whether child and adult readers' comprehension was affected by the use of foreignised items. Using eye movement data and responses to comprehension questions, she found that

the use of foreignising strategies does, in some instances, result in increased cognitive effort compared with domesticated items. She stressed, at various stages, the need for far more evidence of the ways in which translations are received, a view that sits very much at the heart of this paper. Despite the innovative nature of her research, little research can be found since this time to build on this important work and explore other aspects of the ways in which translations are experienced.

Around a similar time to H. Kruger's work, Doherty and colleagues started to conceive of eye tracking as a way to gauge machine translation (MT) quality, based on the hypothesis that increased visual attention would signal higher levels of cognitive effort, and therefore could be correlated with lower levels of user experience (Doherty & O'Brien 2014; Doherty et al. 2010). Doherty et al. (2010) found that poor MT quality tended to result in higher levels of cognitive effort. The later study by Doherty and O'Brien (2014) corroborated these findings, concluding that poor raw MT output yielded higher levels of cognitive effort due to its lower usability.

In a rare example of an experiment involving heart rate monitoring, Rojo et al. (2014) turned their attention to the emotional response of readers to metaphorical and non-metaphorical translations of figurative expressions from English into Spanish. Based on 28 metaphorical expressions portraying happiness, sadness, fear, and rage, they found that the non-metaphorical translations tended to result in a reduced emotional impact, manifested through an increase or decrease in mean heart rate, depending on the type of emotion. As the authors admit, however, the sample size was small, but the importance of this work should not be understated for forging a course for future heart rate-based studies of the reading experience.

In a similar vein to H. Kruger's study, Walker (2018; 2019; 2020) compared the reading experience of stylistic language varieties in literature using a mixed-methods approach involving a preliminary stylistic analysis, and an eye-tracking quasi-experiment. Unlike H. Kruger's comparison of two versions of a translation produced using two different translation strategies, Walker identified and compared the reading experience of stylistically salient elements in the French ST, together with their corresponding translations in the English TT, and a second stylistically 'neutralised' version of the TT, to explore equivalent effect using experimental methods. The data showed that eye-tracking methods can be fruitfully used to compare the cognitive effort experienced at corresponding points in different ST and TT versions of a text.

Finally, a small-scale pilot study was conducted at University College London by Federico Federici (2018a; 2018b) to explore the processing of crucial medical information in crisis contexts as part of the wider INTERACT Crisis Translation project (Crisis Translation n.d.). In this study, Federici combined eye tracking, facial expression analysis, electrocardiography, and electrodermal activity sensors with a questionnaire to investigate how medical professionals understood and responded to three different Italian translations of an original English medical text. One translation was produced by Google Translate, another by Google Translate and post-edited by a native Italian medical professional, and the third was a translation by a non-specialist professional translator. The data exhibited considerable promise for a larger-scale study in future, with participants exhibiting higher levels of cognitive effort and emotional signs of frustration with the less fluent Google Translate output.

To the best of my knowledge, the summary above constitutes the extent of experimental research into the reception of translated texts. When we have such a wealth of available methodological insight and know-how and research output in the fields of natural reading research, cognitive psychology, and the empirical study of literature, it seems as though TS was not fully taking advantage of the available opportunities and its own unique multidisciplinary position. The Empirical Study of Literature, for instance, championed most ardently by Sanford and Emmott in their stimulating book *Mind, Brain and Narrative* (2012), may provide a strong foundation for psychological and neuroscientific work in TS on the different ways in which authors manipulate linguistic features to influence the reading experience. Indeed, the latter part of this paper's subtitle – *Experimental Studies of Translation Reception* – is a subtle nod to this area of research. The intended focus is not the act of producing the TT, but rather the process of experiencing and receiving translations as textual artefacts.

Reception or Experience? (Or Both?)

So what is it that we should be trying to observe, and how should we define the object of our interest? For the remainder of this paper, I will limit discussion to the topic of literary translation, which seems to have yielded a large amount of scholarship in TS, but has been subject to very little (quasi-)experimental investigation, especially on the reception of literary translations. That notwithstanding, I intend my discussion to be more widely applicable to other text types and genres, as my aim is to see these sorts of experimental methods applied to any and all forms of translated texts.

According to various literary traditions, including the reader-response and New Criticism movements, how readers receive a text has generally been extracted from features in the text itself, based on the analysis and intuitions of the critic. Such critics did so by focusing on an 'implied reader' (Iser 1974), among various other related terms, and trying to forge connections between linguistic and stylistic features and the literary effects of the choices on the reader. Other, more ardent reader-response critics, such as Stanley Fish, turned towards trying to establish collective interpretations of texts by a wider readership, through his notion of interpretive communities, based on two principles: '(1) The same reader will perform differently when reading two 'different' […] texts; and (2) different readers will perform similarly when reading the 'same' […] text' (Fish 1976: 481). The difficulty posed by these traditional approaches is that they are based on an abstraction of what is *assumed* to be the reaction of a reader to a particular text or a specific point in a text. The very same criticism can be levelled at conventional comparative analyses of STs and TTs in translation. In recent years, however, there have been a few attempts to develop a more objective approach to exploring the reception of real users of translation.

Conceptually, 'reception' takes on various definitions depending on a range of factors (some intrinsic to specific disciplines, and some more closely linked to different media). I have chosen to adopt Gambier's definition of reception, posited in the context of audiovisual translation, in part because it is a relatively clearexplanation of what reception might comprise, but also because it is paradigmatically broad and can be

applied to and adjusted to account for a variety of different contexts. Gambier contrasts 'reception' with 'perception', defining these terms as follows:

> Perception could be defined as what is impressed on the eyes when watching a film and the way in which viewers represent the viewing act: how they think they watch a film, how they believe they apprehend the viewing process. Perception is made of opinions and impressions and varies over time. Studying reception means to investigate the way(s) in which AV *products/performances are processed, consumed, absorbed, accepted, appreciated, interpreted, understood and remembered* by the viewers, under specific contextual/socio-cultural conditions and with their memories of their experience as cinema going. […] In other words, reception studies in AVT seek to describe and explain what viewers do with the AVT products they are watching or that they have watched. (Gambier 2018: 56, emphasis added)

He further delineates three types of reception, namely response ('perceptual decoding'), reaction ('readability'), and repercussion ('an attitudinal issue' and the 'sociocultural dimension') (Ibid.: 57), all of which relate to various observations of 'events' and/or subsequent reporting or inference of 'acts' (employed, by analogy here, with reference to Toury 1995/2012: 67-68). In more traditionally empirical designs (for instance, Leppihalme 1997; Puurtinen 1995), 'reception' has typically been probed by means of questionnaires, interview data, or think-aloud protocols alone, which are necessarily after the fact and are also reliant on participants' subjective perceptions (Jakobsen & Alves 2020: 4).

This observation also calls to mind Krings' famous diagram of data analysis methods (2005: 348), in which he refers to offline methods ('offline-verfahren') and online methods ('online-verfahren'). Under this typological framework, traditional ST-TT analyses, retrospection, interviews, questionnaires and other comparable methods all fall under the offline category as they relate to participants relaying their thoughts on a particular product after the reading/viewing process and sometimes 'in retrospect whilst watching replays of the completed task' (Hansen-Schirra & Gutermuth 2015: 64). I would like to propose a distinction, therefore, between the term 'reception' and the term 'experience'. 'Reception', I would argue, should be restricted more to the *post hoc* act: the user reads a text, 'responds' in some capacity, and this response is collected using offline methods. 'Experience', in contrast, is more consistent with Krings' online methods of research. Online methods – which Krings defined as encompassing tools such as video recording, observations, eye tracking, EEG, PET, and other similar experimental tools – allow for the moment-to-moment, *in actu* impressions of a reader to be investigated and charted live as the reader progresses through a text (the reader's 'experience').

This distinction has been touched upon in media studies, but has received relatively little attention. Gentikow, for example, argued that the term 'reception' is merely a weak metaphorical extrapolation of the phenomenon, proposing instead that we consider the use of different media in terms of 'experiences':

> The term experience articulates our physical presence in the world. Experience encompasses practical encounters with facts and events of the world, physical and perceptual contact with people and things. Experiences are made primarily by our bodies and senses, are processed cognitively, are learned of, and result in skills, knowledge and values. (Gentikow 2005; in Ytre-Arne 2011: 469)

Ytre-Arne (paraphrasing Gentikow) asserts that 'experiences are made personally, but within and in interaction with society and culture. [They] can be products of active decisions, [but] can also be experienced more or less voluntarily, more or less consciously, as a backdrop in everyday life' (Ibid.). Phenomenologists such as Edmund Husserl would of course have placed considerable emphasis on the role of the body in how we experience the world, and Ytre-Arne picks up on this theme (2011: 474ff), pointing to other phenomenologists such as Simone De Beauvoir, Maurice Merleau-Ponty, and Don Ihde, and stressing that bodily experiences are important in terms of how we interact with media (in the broadest of senses).

In literary studies, Castiglione (2017) laments the lack of empirical (and indeed experimental) research on how stylistic features affect readers, and specifically distinguishes between 'offline' and 'online difficulty' when reading poetry. He defines offline difficulty as 'a post-reading feeling of incomplete or unsatisfactory understanding […] related to the "elusiveness" of poetic meaning', while online difficulty is 'the real-time effort experienced when reading a poem, and is therefore associated with textual resistance' (Ibid.:104). Hence, using my conceptualisation of 'reception', the offline methods proposed by Krings offer a means to capture Castiglione's offline difficulty by probing this 'post-reading feeling' (Castiglione 2017) after the fact; 'experience' would require Krings' online methods to access the 'real-time effort' (Castiglione 2017) involved.

In the context of literary texts, which are my focus in this paper, it is not enough to talk about 'reception' alone, but to consider, more generally the notion of 'experience': what takes place physiologically, emotionally, and cognitively *during* the act of reading, but without discounting the value of perceptions and reflections after the act itself. Despite this seemingly obvious truism, there have been relatively few attempts to investigate this process even though, as Bortolussi and Dixon rightly observe in their book on psychonarratology:

> How readers process narrative is essentially an empirical question that can only be answered by systematic observation of actual readers reading actual texts; it cannot be answered solely on the basis of intuition, anecdotal evidence, or even sophisticated models of human experience. (Bortolussi & Dixon 2003: 13)

This observation therefore begs the question of why more empirical and experimental research has not been conducted to observe 'actual readers reading actual texts'. In translation, therefore, why should we rely on comparative criticism of ST and TT, for example, when so many techniques are now available to explore the ways in which real readers interact with and experience translations?

The Experimental Researcher's Toolkit

For the purposes of this review section, I do not intend to dismiss more traditional, tried and tested offline methods of reception-oriented research such as interviews, questionnaires and retrospection, as well as more basic temporal metrics such as task time and self-paced reading exercises, but rather to focus more on different, and in some cases more innovative techniques which are becoming increasingly accessible (both methodologically and financially) to researchers in TS with a view to stimulating an uptake in the use of

such tools. A good introductory primer on research methodologies in TS, covering the aforementioned offline methods, *inter alia*, can be found in Saldanha and O'Brien (2014). Despite their absence from this section, I maintain that such methods play an important role with regard to triangulation of results with online methods, and in particular to offer qualitative nuance to complement (or indeed contrast with) the quantitative data yielded by the online methods that follow.

Eye tracking and pupillometry

The point-of-gaze paradigm in eye tracking has been used most extensively in empirical studies in TS on account of the widely-accepted view that there is a strong link between the properties of written language and eye movements. The 'linguistic/cognitive position' in natural reading research (for an overview, see Rayner & Liversedge 2011) holds that the duration of time that the eyes linger over a word (a fixation) can be broadly correlated with the amount of cognitive effort involved in processing the fixated word. This brief summary is a vast oversimplification of a huge body of eye-tracking research, as the link between fixations and processing is not quite as direct as the oft-cited 'eye-mind assumption' (Just & Carpenter 1980) would suggest. This observation does not invalidate the data obtained using eye tracking; rather, the data need to be interpreted with a degree of caution and awareness of alternative causes of longer fixations. Nonetheless, there is still tremendous scope for the use of eye tracking in reception-oriented research, in particular to investigate the cognitive dimensions of the reading process on a word, phrase, sentence or higher level, using authentic texts and experimentally manipulated texts, but also as a method to triangulate other data with specific stimuli. See Hvelplund (2017) and Walker and Federici (2018) for overviews of how eye tracking has been used across various paradigms of CTS.

A related approach involving eye tracking is pupillometry, which was first used in TS in O'Brien's (2006) study on the processing of translation memory matches. Pupil dilation is generally considered to change in proportion to mental workload, with pupil diameter growing as mental processing increases. However, in addition to these cognitive dimensions, pupil diameter has also been shown to increase in response to emotional arousal, anticipation, and curiosity (for a summary, see Holmqvist et al. 2015: 393-394). Pupillometry has not been used nearly as extensively in TS as the point-of-gaze paradigm, but, with careful experimental designs, pupil diameter data may prove to be a useful tool, especially when triangulated with other data, in understanding broader emotional responses to textual stimuli, especially in a literary context. Pupillometry is not without its difficulties, however. There is a recognised lag (in some cases, as much as 1,000 ms) between the onset of a stimulus and the reaction of the pupil, which needs to be borne in mind when interpreting the data. The data generated by pupillometry is also extremely 'noisy', and requires careful cleaning before analysis. Pupil size is also influenced by other factors beyond the stimulus itself, and can be affected by caffeine, fatigue, room lighting, and changes in stimulus (screen) brightness, as well as varying in terms of the scale of response between different subjects. In short, pupillometry is a useful tool, but the data needs to be handled with considerable care (as is arguably equally true of many of these tools). A good introduction to pupillometry and its use in other areas of research can be found in Laeng et al. (2012) and Holmqvist et al. (2015: 391-394).

Facial expression analysis (FEA)

Contemporary automated facial expression analysis is founded on a taxonomy of facial movements codified by Ekman and Friesen (1978), and later updated by Ekman et al. (2002), called Facial Action Coding System (FACS). This system of 46 observable action units uses movements of specific parts of the face (inner brow raiser, nose wrinkle, lip pressor, etc.), head movements (head turn left, head tilt right, head shake side to side, etc.), eye movements (eyes turn left, eyes up, cross-eye, etc.), and so-called 'gross behaviour codes' (lip bite, sniff, wink, etc.)[1] to compute not only basic emotions, but also differing degrees of 'valence' and 'arousal'. There are two schools of thought surrounding emotions: on the one hand, some believe that there are six discrete emotions: happiness, sadness, fear, surprise, disgust, and neutral (of which all other emotions are derivatives), but others support a dimensional approach, plotting valence (positive versus negative emotions) against arousal (activating versus calming emotions), allowing for a more complex array of subtle emotions such as nervous, bored, excited, and relaxed (see Selvaraj et al. 2013). FEA studies are remarkably simple to conduct, requiring only a computer webcam and appropriate analysis software. FEA software – such as FaceReader (Noldus, n.d.) and Affectiva Affdex (iMotions, n.d.) – handles the analysis of these various facial movements and can express not only the emotion detected, but also the strength of the emotion. The software has also been found to show a high degree of accuracy in identifying facial expressions of emotion compared with more traditional approaches such as electromyography (see below) (Kulke et al. 2020; Stöckli et al. 2018). The possibilities offered by FEA could provide fruitful results in literary reception research, especially when triangulated with other qualitative (self-reported perceptions) and quantitative data (other biometric measurements). However, as noted by Dupré et al. (2020), just because a person appears to be smiling, it does not mean that he or she is happy. Data triangulation is therefore essential, to gather as much additional data as possible on the emotional and cognitive state of participants, to mitigate the risk of false perceptions of emotions.

Facial electromyography (fEMG)

Facial electromyography is closely related to FEA, but instead of relying on software interpretations of a video input, fEMG measures muscle movements by detecting the electrical activity associated with muscle contractions. fEMG studies require the placement of an array of electrodes on the surface of the skin in positions specific to the facial area of interest. For example, placement of the electrodes at points around the corrugator supercilii muscle (which draws the eyebrow downward and towards the centre of the face, producing a frown) tends to be associated with negative (Bayer et al. 2010) or cognitively-demanding stimuli (Schacht et al. 2010). In contrast, the zygomaticus mayor muscle draws the corners of the mouth upwards and outwards, typically yielding a smile (Lang et al. 1998). By combining fEMG measurements at these two sites, negative and positive emotions can be distinguished with some degree of reliability (for a brief literature review, see Kulke et al. 2020). fEMG offers greater confidence in the detected facial movements given that they are based on direct measurement of muscle contractions, but the range of detectable emotions

is far more restricted than FEA, which, despite being reliant on software interpretation of the visual appearance of the face, allows a wider range of emotions to be detected.

Electrodermal activity (EDA)

Electrodermal activity (also referred to as galvanic skin response) changes in response to external stimuli due to micro-level changes in eccrine sweat gland activity. EDA sensors placed on the hands can detect these minute shifts in perspiration levels caused by emotional arousal. The EDA signal is made up of two elements: skin conductance level, which gradually decreases while participants are at rest and increases rapidly when a new stimulus is presented) and skin conductance response, which exhibits spikes in response to a stimulus (Brishtel et al. 2020). In broad terms, EDA measures the 'physical and psychological state of being alert and ready to act, which can be related to emotional stimulation, increased mental workload, and the startle reflex' (Hartung et al. 2016). While this approach has previously been used in interpreting studies (for instance, Korpal & Jasielska 2019), it has not been used to explore the experience of translations and their source texts (for an excellent critical evaluation of and introduction to using skin conductance methods, see Rojo López & Korpal 2020).

Electrocardiography (ECG)

Electrocardiography can provide insightful data on an individual's physiological or psychological stress levels, by measuring heart rate variability (HRV). It has also been shown, in complex experimental and statistical designs, that it is possible to identify six specific emotional states – happiness, sadness, fear, surprise, disgust and neutral – from ECG data (for example, Selvaraj et al. 2013). In a reading context, therefore, with the right experimental design, it may be possible to observe signs of emotional responses to literary devices such as suspense or linguistic features such as taboo language through an increase on the baseline heart rate (as above, Rojo López & Korpal 2020 provides a thorough review of heart rate and HRV-related approaches).

Electroencephalography (EEG)

Electroencephalography involves the placement of electrodes on the scalp, most frequently using a form of specially designed 'headset' which aligns varying numbers of electrodes with the relevant parts of the brain. The high temporal resolution of EEG data is both a boon and a curse, for it can provide highly detailed insights into the electrical activity of the brain down to the millisecond, but it also produces a vast amount of highly complex data. Oscillations within the theta band (4-8 Hz) are typically associated with task difficulty and working memory; the alpha band (8-12 Hz) is associated with mental attention and alertness; and the beta band (12-25 Hz) is linked to active concentration (Niedermeyer & Lopes da Silva 2012). EEG data is therefore a prime choice for detailed analysis of attention, cognitive load and working memory in different tasks or with different stimuli (see, for instance, Baceviciute et al. 2020). Like pupillometry above, the data is frequently noisy and replete with artefacts, rendering careful data cleaning a necessity. The choice of equipment is also an important factor in improving data collection.

Others

Needless to say, the tools outlined above do not constitute the full array of devices that could be used in experimental studies. Indeed, other techniques such as magnetoencephalography (MEG), positron emission tomography (PET), and functional magnetic resonance imaging (fMRI) exist and have been used to a small degree in cognitive (and neurological) studies of the translation and interpreting process more generally (see García et al. 2016), but these tools are unwieldy and, for most researchers, inaccessible, both financially and methodologically. The various sensors outlined in the previous sub-sections have dropped dramatically in price over the last decade, and now, a full suite of multi-sensory devices can be procured relatively inexpensively. However, it is important that we do not equate accessibility in terms of cost with accessibility in terms of experimental design and the interpretation of data (see Concluding Remarks).

Example: The Pevearsion of Russian Literature

Having outlined some of the methodological approaches that could be used to investigate different aspects of the reading experience, from a cognitive and emotional perspective, we will now consider an example of the ways in which text-based translations could be explored experimentally, based on a comparison between different versions of the TT. The chosen example examines the effects of different translation approaches adopted by different translators, in this case contrasting the award-winning, but problematic translations by Richard Pevear and Larissa Volokhonsky[2] of Russian classics with an older translation by Rosemary Edmonds. The title of this sub-section is borrowed from an article by Morson (2010) strongly criticising Pevear and Volokhonsky's translations for their perversion ('Pevearsion') of Russian classics into 'awkward and unsightly muddles'. Despite their popular acclaim and marketing triumphs, Pevear and Volokhonsky's translations have been criticised by many for 'taking everything they can get their hands on written in Russian and putting it into flat, awkward English' (Malcolm 2016). This stilted style is more than likely a symptom of their unusual translation method, whereby Volokhonsky (a native Russian speaker, with only moderate knowledge of English) renders the Russian in a highly literal English, and Pevear (a native English speaker, who has little knowledge of Russian) then adjusts the literal translation into a purportedly more fluid English translation, resulting in a translation that follows the Russian syntax and phrasing extremely closely (Morson 2010). Consider the two extracts from the opening to Tolstoy's novella *The Death of Ivan Ilyich* below. The first is the 1960 translation by Rosemary Edmonds, the second the 2010 translation by Pevear and Volokhonsky:

> In the great building of the Law Courts, during an interval in the hearing of the Melvinsky affair, the members of the Court and the public prosecutor gathered together in Ivan Yegorovich Shebek's private room, and the conversation turned on the celebrated Krasovsky case. Fiodr Vassilyevich maintained hotly that it was not subject to their jurisdiction, Ivan Yegorovich argued the contrary, while Piotr Ivanovich, not having entered into the discussion at the start, was taking no part in it but looking through the *Gazette* which had just been brought in. (Edmonds' translation; Tolstoy 1960)

> In the big building of the law courts, during a break in hearing the case of the Melvinskys, the members and the prosecutor met in Ivan Yegorovich Shebek's office, and the conversation turned to the famous Krasovsky case. Fyodor Vassilievich became heated demonstrating non-jurisdiction, Ivan Yegorovich stood his ground; as for Pyotr Ivanovich, not having entered into the argument in the beginning, he took no part in it and was looking through the just-delivered *Gazette*. (Pevear and Volokhonsky's translation; Tolstoy 2010)

Edmonds' translation would likely be considered by many as a relatively 'fluent' translation, while Pevear and Volokhonsky's would be seen as 'resistant'. Pevear and Volokhonsky are at pains to retain the 'Russianness' of the original texts and have previously noted that, unlike other translators, they purposely avoid 'smoothing over' the Russian (Remnick 2005). Looking at the extracts above, phrases in Pevear and Volokhonsky's translation such as 'the case of the Melvinskys', 'demonstrating non-jurisdiction', 'not having entered into the argument in the beginning', and 'just-delivered *Gazette*' are particularly jarring and are almost identical in style and syntax to the original Russian[3]. Without digressing into the merits and shortcomings of Pevear and Volokhonsky's style, the comments made by the aforementioned reviewers and others are perfectly amendable to experimental investigation. In the sub-sections that follow, I will delineate some research questions and explain a possible experimental design to investigate these questions.

Research Questions

In essence, what this experimental design is looking to investigate is whether Pevear and Volokhonsky's style is indeed jarring and disruptive to reading compared with Rosemary Edmonds' translation, as argued by its critics. Three research questions to approach this overarching objective might be:

> RQ1: Is there a perceptible difference between the reading experience of Pevear and Volokhonsky's foreignising translation of *The Death of Ivan Ilyich* and Rosemary Edmonds' domesticating translation?
>
> RQ2: To what extent does the cognitive effort of readers differ over parallel foreignised and domesticated elements in the Pevear and Volokhonsky translation and Edmonds translation respectively?
>
> RQ3: How do the sensory data on physiological and emotional responses to the reading experience correlate with areas of heightened or reduced cognitive effort?

The working hypotheses underpinning these questions are that the data will demonstrate a difference in the reading experience of the two texts, in particular in the form of higher levels of cognitive effort over foreignised elements of Pevear and Volokhonsky's translation (manifested through longer fixation durations), but also, potentially, in the form of sensory data pointing to heightened stress levels and frustration at the more laborious reading experience when encountering foreignised elements.

Methods

Because this experiment would be exploratory, it would look to combine a more established experimental method – eye-tracking – with lesser-explored methods such as FEA and EDA. It is also widely recognised that a mixed-methods design combining qualitative and quantitative methods yields a more detailed, multi-

faceted understanding of a topic and can help to improve validity (McKim 2017). A schema of the overall research design is provided in Figure 1.

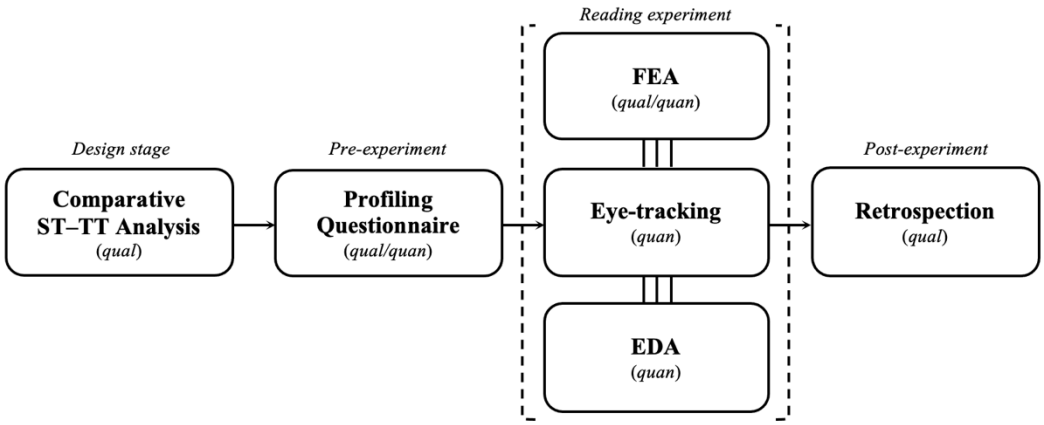

Figure 1. Schema of research design

In a similar procedure to the proof-of-concept presented in Walker (2020), an important first step in terms of operationalising the notions of foreignised and domesticated elements in each translation would be to undertake a comparative analysis of the ST and two TTs with a view to coding, qualitatively, those elements that differ between the translations and whether they are characteristic of a foreignising and domesticating approach. H. Kruger (2016) provides a clear explanation of how fluency and resistancy (lower-level textual features which influence the cognitive processing of readers) relate to domestication and foreignisation (the higher-level strategies or attitudinal principles which affect cognition on a wider scale). Using these definitions, and descriptors relating to notions of fluency and resistancy in the coding of the data, this preliminary qualitative analysis would enable some tentative hypotheses to be raised about areas of the Pevear and Volokhonsky translation where cognitive effort might be higher due to potential syntactic or lexical obfuscation arising from their foreignisation strategy. Being qualitative, and critic-driven, this is an inherently intuitive, and indeed subjective, process, but as argued in Walker (2020: 384ff, inter alia), can still serve an important 'check and balance' role to provide nuance to the experimental data or to explain any unexpected data arising from the experiment. This preparatory qualitative phase could even be complemented by some survey or focus group approaches to ask independent participants to identify aspects of the translations that are particularly difficult to read or are stylistically awkward (i.e. 'resistant'). If low-frequency language is at play in other studies (which has been consistently proven to result in longer fixations when reading; see Cop et al. 2015), corpus-based frequency data can also be integrated in the early stages of the design for correlation with the experimental data.

The experiment would naturally begin with the standard informed consent process and the completion of a profiling questionnaire which can be used to obtain additional information on the participants' backgrounds, reading habits, and any other data that may be useful to categorise the experiment data. The experiment itself would in fact be classified as a quasi-experiment, in this case, as the two texts from which the stimuli are extracted are authentic, unedited translations. A quasi-experiment is:

> any research method that has some of the features of an experiment but is not strictly experimental inasmuch as the investigator either does not manipulate the independent variable directly or does not have full control over the extraneous or nuisance variables that might influence the results (Colman 2008).

There are obvious pros and cons to quasi-experiments. Downsides to quasi-experiments include the lack of experimental control over variables such as ordering effects, and the increased difficulty in generalising causal effects. They can however achieve greater ecological validity, especially in natural reading studies, and can also offer more opportunities for post-hoc analysis of data from different (and sometimes unexpected) perspectives.

The vital linchpin of this proposed experimental design is eye-tracking. During the reading task, participants would read a series of extracts from the Pevear and Volokhonsky translation or the Edmonds translation (or indeed both, as part of a within-subjects design aimed at mitigating habituation). The eye-tracking data will be able to provide a clear record of the participants' reading process as they progress through the text, focus on specific words or phrases, regress to problematic areas, and skim or skip others entirely. This precise temporal and visual record (subject to good calibration, of course) can be broken down into different fixation-based metrics such as 'gaze duration' or 'total fixation duration'[4] to give an insight into the different stages of lexical and syntactic processing, and subsequently analysed in conjunction with the FEA and EDA data.

For example, if the eye-tracking data were to point to heightened cognitive effort over the phrase 'just-delivered *Gazette*' (evidenced through longer fixation durations compared with Edmonds' translation), we could look to the EDA data in this area of interest (AOI) to explore whether there is a skin conductance response. Because EDA has not been used in this way before (to the best of my knowledge), this would be a highly exploratory aspect of the design: to consider whether AOIs such as this, where the eye-tracking data would presumably reveal higher levels of cognitive effort (as hypothesised by the preliminary qualitative analysis), do in fact give rise to skin conductance responses on such a small level. It may be that the response is not perceptible at all, or that the response is delayed, or the skin conductance may still declining from a previously heightened level induced by a prior area of interest. We can also couple the EDA data with the FEA data to see if the software algorithms detected positive or negative valence and the intensity of the emotion experienced. A strong (high-intensity) negative valence could potentially point to a feeling of frustration in this AOI, for example[5]. To use my terminological conceptualisation proposed above, the eye-tracking, FEA and EDA data would provide an insight into the reader's *experience* of the relevant text, but not its *reception*. It is at this point that the 'offline' post-experiment retrospection task plays an important role.

As other researchers have done in the past (for example, Huang 2018), eye-tracking gaze data can be replayed to participants immediately after the main experimental task is complete and participants can be asked to comment on the video replay while the audio is recorded and synchronised with the replayed video. The retrospection task could be cued – where the researcher asks guided questions to prompt a targeted response from the participant over a particularly long fixation or series of multiple regressions, for instance – or it could be uncued – where the participant is asked to reconstruct their experience of the text. Arguably,

cued retrospection is likely to yield more useful data, in particular to tease out some of the nuance behind the areas of heightened cognitive effort evidenced in the eye-tracking data, but also because 'untrained' participants may not focus on matters of relevance when discussing their experience. Retrospection should help to provide a further insight into the reader's reception of the text, guided by the replay of their experience while reading. Longer fixations in the replay may be the result of frustration or difficulty when reading a particularly 'resistant' AOI (a negative experience), but it could equally be a form of narrative engagement or affect (a positive experience). Hence, cognitive effort should not be equated with a negative experience automatically: one of the many roles of literature is to promote learning, engagement, and communication between peoples, societies and cultures, and higher cognitive effort may well reflect such processes as opposed to difficulties with comprehension or a stilted reading experience.

Wider implications?

One methodological difficulty that arises from any experiment using a similar design to the one outlined above is how to achieve wider generalisation of the results. As Fish acknowledged in the 1980s in reference to readers' responses to literary texts, there is *similarity* in readers' experiences, but also considerable idiosyncrasy too (Fish 1976; 1980), an assertion also made by Chesterman in his three laws of translation effects. This debate and possible solutions to this problem are addressed in far greater detail in Walker (2020), where the data showed that differences did exist between participants, but 'spikes' in fixation durations tended to congregate over similar locations in the text across different readers. The challenge was more one of 'normalising' the data so that the data collected from the participants could be compared and more general conclusions extrapolated. For instance, in the aforementioned study, it was shown that where stylistic complexity was high, the *diversity* in reader experiences was also high (some participants showed high fixation durations, while others did not); but where stylistic complexity was low, the reading experience was extremely homogeneous: all readers experienced it in a very similar way. These findings do support, to a small extent, the distinction made earlier between 'experience' and 'reception'. The 'experience' was manifested by the values of the dependent variables across the AOIs, each with differing levels of stylistic complexity and – crucially – this experience, once recorded, does not change. That is not to say that the experience would not be different if the same text were re-read later that day, or one week later, or several years later. Rather, that single experience has been documented and, once documented, cannot change. Despite this, the rationale for the experience at different points in the text cannot be gleaned from the eye-tracking data alone. This rationalisation would need to be derived from an investigation of the readers' 'reception' – i.e. their post-hoc verbalisation and meta-awareness of the possible reasons for having experienced the text in the manner recorded by the eye tracker. Such reception can change even in the short period of time between reading and reporting these comments to the researcher, and such comments can be shaped by the phrasing of the questions by the researcher, the prompts given in the retrospection task, and various other factors, including the mere passage of time, allowing the participant to reflect more deeply on the stimuli itself.

In the case of the experiment proposed above, a careful experiment design and statistical analysis would enable researchers to explore the impact of fluent and resistant passages on the readability of literary texts, allowing for more authoritative conclusions to be put forward on the merits or otherwise of Pevear and Volokhonsky's translation, going beyond mere intuition and subjective opinions, however well informed. The use of eye tracking, in particular, also allows for subsequent analysis of a wide array of alternative dependent variables beyond those mentioned above. Because of the richness of eye-tracking data (both a boon and curse), pupillometry data can be extracted and analysed *post-hoc*, and saccade and regression data can shed new light on perspectives derived from the fixation data. However, it is important that we not lose sight of the bigger picture through this focus on micro-level influences. As Tymoczko rightly observed:

> we need to remain aware of the links between micro and macro research: to engage in investigations of translation at the micro level of the brain but to see as well the implications for the macro levels of translations as texts, as mediations between cultures, and as ideological interventions, as well as the implications for many other macro level topics that have flourished in translation studies (Tymoczko 2012).

Hence, such micro-level studies of the phenomenological effects of linguistic choices by literary authors and literary translators (in this example) should not be detached from the wider socio-cultural influences on experience and reception that are debated in other areas of TS (see Pym 2020). Rather, they should work in harmony and inform and complement one another by drawing on the complementary 'bottom-up' and 'top-down' perspectives that each paradigm offers.

We should also remember that, in the early stages of the proposed expansion of such studies, many of these methods would be highly exploratory. New data would be collected from different sources, on different stimuli, and different experimental and data analysis techniques would need to be used and trialled. The process would therefore be a long-term, iterative process of design, test, analyse, evaluate, and repeat, with each study building on the work of previous ones, ideally with researchers sharing best practices and know-how to take this field forward.

Concluding Remarks

As should hopefully have become clear from the example above, the potential for research into the experience of reading translations is vast, not solely in a literary context – as was the focus in this paper –, but also in the healthcare sector, the legal sector, marketing and business, education, and news, each addressing different text types, functions, and audiences. The examples discussed in this paper focus predominantly on negative cognitive effects and emotions (higher levels of cognitive effort, confusion, frustration, etc.), but precisely the same approaches can be applied to more positive forms of cognition and emotion (learning engagement, comprehension and readability, reactions to humour, the processing of metaphorical language, other forms of narrative immersion such as suspense, etc.).

The techniques discussed in this paper are no longer the reserve of vast, well-funded laboratories. All of these sensors and devices have reduced considerably in cost over the last decade, opening up huge potential for complex, exploratory multi-sensory experiments. Unlike monolingual explorations of reading, translation

provides the perfect forum to explore the reading experience. Since many STs already have multiple translations (in different languages, or in the same language, but by different translators), it can obviate the need to artificially manipulate texts for quasi-experimental investigation. Even where existing translations do not exist, by its very nature, the different solutions chosen at different locations in a text can allow for a more organic experimental manipulation of source text structures, stylistic choices, and assumed effects into different forms in the TT, depending on the translation procedures employed and the *skopos* sought. Despite the accessibility of such methods and the ideal placement of TS as a discipline for such research, we should not discount the importance of multi-skilled, interdisciplinary research teams and should endeavour to move away from the tradition of the lone researcher that has dominated in TS in the past. Yet, the accessibility of experimental technology noted above (in particular, from a financial perspective) should not be equated with the accessibility of methods. As such, I would argue that research groups combining the expertise of colleagues from TS, statistics, cognitive psychology, and other cognate fields are essential to navigate the complexity of experimental research and achieve the returns that we so desire. Such a shift in research collaborations will necessarily entail a rethinking of not only the methodological, but also the institutional challenges posed by such interdisciplinarity (see Federici & Walker 2018: 24).

The time is ripe to harness the vast untapped potential in this area and arrive at a better understanding of the effects that different translation strategies yield and the potential variation in effects between source and target text. But such endeavours need to be done right. Researchers would do well to heed the sage advice advocated in Orero et al. (2018) on the subject of conducting experimental research in AVT. (Perhaps a similar position paper is also necessary for the specific characteristics of investigating the experience of text-based translations too?) In short, we need more data. We need data on how people read texts of different genres, in different languages, in different styles. Such data are not solely for the benefit of the research findings alone, but the data will also aid the refinement of best practices in text-based reception studies. Some experiments will be more successful than others: some will yield results entirely incompatible with the posited hypotheses, while others will produce data exactly as hoped. We also need more *sharing* of experimental data too, allowing researchers to re-analyse, re-interpret and explore the data from new perspectives and facilitating meta-analyses of findings across various related studies over time. All of these data are essential to further unpack and disentangle the mystery surrounding the cognitive, emotional, and physiological reactions to different stimuli, triangulating data from multiple sources, and comparing data between ST and TT readers. What is abundantly clear at present is that we simply do not know enough about how readers actually interact with translations on a micro- and macro-level. This situation needs to and *can* change; it could have a profound effect on the way that we conceptualise the process of translation across all genres, and could lead to a re-evaluation of theories and concepts at the very core of translation studies.

References

Alves, F., & Hurtado Albir, A. (2017). Evolution, Challenges, and Perspectives for Research on Cognitive Aspects of Translation. In J. W. Schwieter & A. Ferreira (Eds.), *The Handbook of Translation and Cognition* (pp. 537-554). Malden, MA: John Wiley & Sons, Inc.

Alves, F., & Jakobsen, A. L. (Eds.). (2021). *The Routledge Handbook of Translation and Cognition*. Abingdon and New York: Routledge.

Alves, F., Pagano, A., & da Silva, I. (2011). Towards an investigation of reading modalities in/for translation: An Exploratory study using eye-tracking data. In S. O'Brien (Ed.), *Cognitive Explorations of Translation* (pp. 175-196). London: Continuum.

Baceviciute, S., Mottelson, A., Terkildsen, T., & Makransky, G. (2020). *Investigating Representation of Text and Audio in Educational VR Using Learning Outcomes and EEG*. Paper presented at the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu. 25-30 April 2020. http://aske.mottelson.dk/wp-content/uploads/2020/03/Paper743.pdf

Balling, L. W., & Hvelplund, K. T. (2015). Design and statistics in quantitative translation (process) research. *Translation Spaces, 4*(1), 169-186.

Bayer, M., Sommer, W., & Schacht, A. (2010). Reading Emotional Words Within Sentences: The Impact of Arousal and Valence on Event-Related Potentials. *International Journal of Psychophysiology, 78*, 299-307.

Bortolussi, M., & Dixon, P. (2003). *Psychonarratology: Foundations for the Empirical Study of Literary Response*. Cambridge: Cambridge University Press.

Brems, E., & Ramos Pinto, S. (2013). Reception and translation. In Y. Gambier & L. v. Doorslaer (Eds.), *Handbook of Translation Studies* (Vol. 4, pp. 142-147). Amsterdam: John Benjamins Publishing.

Brishtel, I., Khan, A. A., Schmidt, T., Dingler, T., Ishimaru, S., & Dengel, A. (2020). Mind Wandering in a Multimodal Reading Setting: Behavior Analysis and Automatic Detection Using Eye-Tracking and an EDA Sensor. *Sensors, 20*(9), 2546.

Castiglione, D. (2017). Difficult poetry processing: Reading times and the narrativity hypothesis. *Language and Literature, 26*(2), 99-121.

Chesterman, A. (1998). Causes, Translations, Effects. *Target, 10*(2), 201-230.

Colman, A. M. (2008). *Dictionary of Psychology (3rd Edition)*. Oxford: Oxford University Press.

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency Effects in Monolingual and Bilingual Natural Reading. *Psychonomic Bulletin and Review, 22*(5), 1216–1234.

Crisis Translation. (n.d.). INTERACT International Network on Crisis Translation. Retrieved from https://sites.google.com/view/crisistranslation/consortium (Accessed on 10 July 2020).

d'Ydewalle, G., Van Rensbergen, J., & Pollet, J. (1987). Reading a Message When the Same Message Is Available Auditorily in Another Language: the Case of Subtitling. In J. K. O'Reagan & A. Lévy Schoen (Eds.), *Eye Movements: From Physiology to Cognition* (pp. 313-321). Amsterdam: Elsevier.

Doherty, S., & O'Brien, S. (2014). Assessing the Usability of Raw Machine Translated Output: A User-Centered Study Using Eye Tracking. *International Journal of Human-Computer Interaction, 30*(1), 40-51.

Doherty, S., O'Brien, S., & Carl, M. (2010). Eye Tracking as an MT Evaluation Technique. *Machine Translation, 24*, 1-13.

Dragsted, B., & Hansen, I. G. (2008). Comprehension and production in translation: A Pilot study on segmentation and the coordination of reading and writing processes. In I. M. Mees, F. Alves, & S. Göpferich (Eds.), *Methodology, Technology and Innovation in Translation Process Research* (pp. 9-29). Frederiksberg: Samfundslitteratur.

Dupré, D., Krumhuber, E. G., Küster, D., & McKeown, G. J. (2020). A Performance Comparison of Eight Commercially Available Automatic Classifiers for Facial Affect Recognition. *PloS one, 15*(4), e0231968. doi:10.1371/journal.pone.0231968

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System: The Manual on CD-ROM*. Salt Lake City: A Human Face.

Federici, F. M. (2018a). *A Mapping Exercise: Eye Tracking, Translation, and Multi-sensorial Data*. Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa (INESC-ID), Lisbon, Portugal. 6 April 2018.

Federici, F. M. (2018b). *Tracking Eyes and Chasing Comprehension: From English as Lingua Franca to English As Pivot Language to Access Crucial Information*. Paper presented at the Biennial Conference of the Associazione Italiana di Anglistica [Italian Association of English Studies], University of Macerata, Italy. 18-19 May 2018.

Federici, F. M., & O'Brien, S. (Eds.). (2019). *Translation in Cascading Crises*. Abingdon: Routledge.

Federici, F. M., & Walker, C. (2018). A Mapping Exercise: Eye Tracking and Translation. In C. Walker & F. M. Federici (Eds.), *Eye-tracking and Multidisciplinary Studies on Translation* (pp. 11–29). Amsterdam: John Benjamins.

Fish, S. (1976). Interpreting the "Variorum". *Critical Inquiry, 2*(3), 465-485.

Fish, S. (1980). *Is There A Text In This Class? The Authority of Interpretive Communities*. Cambridge: Harvard University Press.

Gambier, Y. (2018). Translation Studies, Audiovisual Translation and Reception. In E. Di Giovanni & Y. Gambier (Eds.), *Reception Studies and Audiovisual Translation* (pp. 43-67). Amsterdam: John Benjamins.

García, A. M., Mikulan, E., & Ibáñez, A. (2016). A Neuroscientific Toolkit for Translation Studies. In R. Muñoz Martín (Ed.), *Reembedding Translation Process Research* (pp. 21-46). Amsterdam: John Benjamins.

Gentikow, B. (2005). *Exploring Media Experiences: A New Approach to Reception Theory and Empirical Studies*. Paper presented at the First European Communication Conference, Amsterdam. November 2005.

Hansen-Schirra, S., & Gutermuth, S. (2015). Approaching comprehensibility in translation studies. In K. Maksymski, S. Gutermuth, & S. Hansen-Schirra (Eds.), *Translation and Comprehensibility* (pp. 53-77). Berlin: Frank & Timme.

Hartung, F., Burke, M., Hagoort, P., & Willems, R. M. (2016). Taking Perspective: Personal Pronouns Affect Experiential Aspects of Literary Reading. *PloS one, 11*(5), e0154732. doi:10.1371/journal.pone.0154732

Hoban, R. (2002 [1980]). *Riddley Walker*. London: Bloomsbury Publishing.

Hodson, J. (2014). *Dialect in film and literature*. Basingstoke: Palgrave Macmillan.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van De Weijer, J. (2015). *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford: Oxford University Press.

Holz-Mänttäri, J. (1984). *Translatorisches Handeln: Theorie und Methode*. Helsinki: Suomalainen Tiedeakademia.

Huang, J. (2018). Working Style of Student Translators in Self-Revision, Other-Revision and Post-Editing. In C. Walker & F. M. Federici (Eds.), *Eye Tracking and Multidisciplinary Studies on Translation* (pp. 145-184). Amsterdam: John Benjamins.

Hvelplund, K. T. (2017). Eye Tracking in Translation Process Research. In J. W. Schwieter & A. Ferreira (Eds.), *The Handbook of Translation and Cognition* (pp. 248-264). Malden, MA: John Wiley & Sons, Inc.

iMotions. (n.d.). Affectiva iMotions Biometric Research Platform. Retrieved from https://imotions.com/affectiva/ (Accessed on 9 July 2020).

International Organization for Standardization. (2018). Ergonomics of Human-System Interaction – Part 11: Usability: Definitions and Concepts (ISO 9241-11:2018). Geneva: International Organization for Standardization.

Iser, W. (1974). *The Implied Reader: Patterns of Communication in Prose, Fiction from Bunyan to Beckett*. Baltimore: John Hopkins University Press.

Jakobsen, A. L. (2017). Translation Process Research. In J. W. Schwieter & A. Ferreira (Eds.), *The Handbook of Translation and Cognition* (pp. 21-49). Maiden, MA: John Wiley & Sons, Inc.

Jakobsen, A. L., & Alves, F. (2020). Introduction. In F. Alves & A. L. Jakobsen (Eds.), *The Routledge Handbook of Translation and Cognition* (pp. 1-22). Abingdon and New York: Routledge.

Jakobsen, A. L., & Jensen, K. T. H. (2008). Eye Movement Behaviour Across Four Different Types of Reading Task. In S. Göpferich, A. L. Jakobsen, & I. M. Mees (Eds.), *Looking at Eyes: Eye-Tracking Studies of Reading and Translation Processing.* (pp. 103-124). Frederiksberg: Samfundslitteratur Press.

Jensen, K. T. H. (2011). Distribution of Attention Between Source Text and Target Text During Translation. In S. O'Brien (Ed.), *Cognitive Explorations of Translation* (pp. 215-237). London: Continuum.

Jordison, S. (2017). Worth the 'Trubba': Making Sense of Riddley Walker's Language. *The Guardian*. Retrieved from https://www.theguardian.com/books/booksblog/2017/nov/07/worth-the-trubba-making-sense-of-riddley-walkers-language (Accessed on 10 July 2020).

Just, M. A., & Carpenter, P. A. (1980). A Theory of Reading: From Eye Fixations to Comprehension. *Psychological Review, 87*(4), 329-354.

Korpal, P., & Jasielska, A. (2019). Investigating Interpreters' Empathy: Are Emotions in Simultaneous Interpreting Contagious? *Target, 31*(1), 2-24.

Kotze, H. (2020). Converging *What* and *How* to Find Out *Why*: An Outlook on Empirical Translation Studies. In L. Vandevoorde, J. Daems, & B. Defrancq (Eds.), *New Empirical Perspectives on Translation and Interpreting* (pp. 333-371). London: Routledge.

Krings, H. P. (2005). Wege ins Labyrinth – Fragestellungen und Methoden der Übersetzungsprozessforschung im Überblick. *Meta, 50*(2), 345-358.

Kruger, H. (2012). *Postcolonial polysystems: the production and reception of translated children's literature in South Africa*. Amsterdam: John Benjamins.

Kruger, H. (2013). Child and adult readers' processing of foreignised elements in translated South African picturebooks: an eye-tracking study. *Target, 25*(2), 180-227.

Kruger, H. (2016). Fluency/resistancy and domestication/foreignisation. *Target, 28*(1), 4-41.

Kruger, H., & Kruger, J.-L. (2017). Cognition and reception. In J. W. Schwieter & A. Ferreira (Eds.), *The Handbook of Translation and Cognition* (pp. 71-89). Malden, MA: John Wiley & Sons.

Kruger, J.-L. (2018). Eye Tracking in Audiovisual Translation Research. In L. Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation* (pp. 350-366). London: Routledge.

Kruger, J.-L., Hefer, E., & Matthew, G. (2013). *Measuring the impact of subtitles on cognitive load: Eye tracking and dynamic audiovisual texts.* Paper presented at the Proceedings of the 2013 Conference on Eye Tracking South Africa.

Kulke, L., Feyerabend, D., & Schacht, A. (2020). A Comparison of the Affectiva iMotions Facial Expression Analysis Software With EMG for Identifying Facial Expression of Emotion. *Frontiers in Psychology, 11*(329). doi:10.3389/fpsyg.2020.00329

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A Window to the Preconscious? *Perspectives on Psychological Science, 7*(1), 18-27.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, Motivation, and Anxiety: Brain Mechanisms and Psychophysiology. *Biological Psychiatry, 44*, 1248-1263.

Leppihalme, R. (1997). *Culture Bumps: An Empirical Approach to the Translation of Allusions*. Clevedon: Multilingual Matters.

Malcolm, J. (2016). Socks. *The New York Review of Books*. Retrieved from https://www.nybooks.com/articles/2016/06/23/socks-translating-anna-karenina/ (Accessed on 10 July 2020).

McKim, C. A. (2017). The Value of Mixed Methods Research: A Mixed Methods Study. *Journal of Mixed Methods Research, 11*(2), 202-222.

Morson, G. S. (2010). The Pevearsion of Russian Literature. *Commentary, 130*(1), 92-92.

Muñoz Martín, R. (2017). Looking Toward the Future of Cognitive Translation Studies. In J. W. Schwieter & A. Ferreira (Eds.), *The Handbook of Translation and Cognition* (pp. 555-572). Malden, MA: John Wiley & Sons, Inc.

Nida, E. (1964). *Toward a science of translating with special reference to principles and procedures involved in Bible translating*. Leiden: E. J. Brill.

Niedermeyer, E., & Lopes da Silva, F. H. (2012). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields* (6th ed.). Philadelphia: Lippincott Williams and Wilkins.

Noldus. (n.d.). FaceReader. Retrieved from https://www.noldus.com/facereader (Accessed on 9 July 2020).

Nord, C. (2005). *Text Analysis in Translation* (C. Nord & P. Sparrow, Trans.). Amsterdam: Rodopi.

O'Brien, S. (2006). Eye-tracking and Translation Memory Matches. *Perspectives, 14*(3), 185-205.

O'Brien, S. (2009). Eye Tracking in Translation Process Research: Methodological Challenges and Solutions. In I. M. Mees, F. Alves, & S. Göpferich (Eds.), *Methodology, Technology and Innovation in Translation Process Research*. Copenhagen: Samfundslitteratur.

O'Brien, S. (2013). The Borrowers: Researching the Cognitive Aspects of Translation. *Target, 25*(1), 5-17.

O'Brien, S. (Ed.) (2011). *Cognitive Explorations of Translation*. London: Continuum.

Olalla-Soler, C., Franco Aixelá, J., & Rovira-Esteva, S. (2020). Mapping Cognitive Translation and Interpreting Studies: A Bibliometric Approach. *Linguistica Antverpiensia, New Series: Themes in Translation Studies, 19*, 25-52.

Orero, P., Doherty, S., Kruger, J.-L., Matamala, A., Pedersen, J., Perego, E., . . . Szarkowska, A. (2018). Conducting Experimental Research in Audiovisual Translation (AVT): A Position Paper. *The Journal of Specialised Translation, 30*, 105-126.

Orrego-Carmona, D. (2018). Audiovisual Translation and Audience Reception. In L. Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation* (pp. 367-382). London: Routledge.

Porter, J. (1990). "Three Quarks for Muster Mark": Quantum Wordplay and Nuclear Discourse in Russian Hoban's "Riddley Walker". *Contemporary Literature, 31*(4), 448-469.

Puurtinen, T. (1995). *Linguistic Acceptability in Translated Children's Literature.* (PhD Thesis), University of Joensuu, Joensuun Yliopiston Humanistisia Julkaisuja, 15.

Pym, A. (2020). *How Are Translations Really Received?* Paper presented at the Cardiff University School of Modern Languages Research Seminars, Cardiff University. 19 November 2020. https://www.youtube.com/watch?v=FjKbGEWQECc

Rayner, K., & Liversedge, S. P. (2011). Linguistic and cognitive influences on eye movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 751-764). Oxford: Oxford University Press.

Reiss, K., & Vermeer, H. J. (1984). *Grundlegung einer allgemeinen Translationstheorie [Groundwork for a General Theory of Translation]*. Tübingen: Niemeyer.

Remnick, D. (2005). The Translation Wars. *The New Yorker*. Retrieved from https://www.newyorker.com/magazine/2005/11/07/the-translation-wars (Accessed on 10 July 2020).

Rojo, A., Ramos, M., & Valenzuela, J. (2014). The Emotional Impact of Translation: A Heart Rate Study. *Journal of Pragmatics, 71*, 31-34.

Rojo López, A. M., & Korpal, P. (2020). Through Your Skin to Your Heart and Brain: A Critical Evaluation of Physiological Methods in Cognitive Translation and Interpreting Studies. *Linguistica Antverpiensia, New Series: Themes in Translation Studies, 19*, 191-217.

Saldanha, G., & O'Brien, S. (2014). *Research Methodologies in Translation Studies*. London: Routledge.

Sanford, A. J., & Emmott, C. (2012). *Mind, Brain and Narrative*: Cambridge University Press.

Schacht, A., Dimigen, O., & Sommer, W. (2010). Emotions in Cognitive Conflicts Are Not Aversive But Are Task Specific. *Cognitive, Affective and Behavioral Neuroscience, 10*(3), 349-356.

Schaeffer, M., Nitzke, J., & Hansen-Schirra, S. (2019). Predictive Turn in Translation Studies: Review and Prospects. In S. Brunn & R. Kehrein (Eds.), *Handbook of the Changing World Language Map* (pp. 3939–3961). Cham: Springer.

Schwieter, J. W., & Ferreira, A. (Eds.). (2017). *The Handbook of Translation and Cognition*. Malden, MA: John Wiley & Sons, Inc.

Selvaraj, J., Murugappan, M., Wan, K., & Yaacob, S. (2013). Classification of Emotional States from Electrocardiogram Signals: A Non-Linear Approach Based on Hurst. *BioMedical Engineering OnLine, 12*(44).

Sjørop, A. C. (2008). Metaphor Comprehension in Translation: Methodological Issues in a Pilot Study. In S. Göpferich, A. L. Jakobsen, & I. M. Mees (Eds.), *Looking at Eyes: Eye-Tracking Studies of Reading and Translation Processing.* (pp. 53-78). Frederiksberg: Samfundslitteratur Press.

Sjørop, A. C. (2011). Cognitive effort in metaphor translation: An Eye-tracking study. In S. O'Brien (Ed.), *Cognitive Explorations of Translation* (pp. 197-214). London: Continuum.

Sjørop, A. C. (2013). *Cognitive effort in metaphor translation.* (PhD Dissertation), Samfundslitteratur, Copenhagen.

Snell-Hornby, M. (1995). *Translation Studies: An Integrated Approach* (2nd ed.). Amsterdam: John Benjamins.

Stöckli, S., Schulte-Mecklenbeck, M., Borer, S., & Samson, A. C. (2018). Facial Expression Analysis with AFFDEX and FACET. *Behavior Research Methods, 50*, 1446-1460. doi:10.3758/s13428-017-0996-1

Suojanen, T., Koskinen, K., & Tuominen, T. (2015). *User-Centered Translation*. Abingdon: Routledge.

Teixeira, C. S. C., & O'Brien, S. (2018). Overcoming Methodological Challenges of Eye Tracking in the Translation Workplace. In C. Walker & F. M. Federici (Eds.), *Eye-tracking and Multidisciplinary Studies on Translation* (pp. 33–54). Amsterdam: John Benjamins.

Tolstoy, L. (1960). *The Death of Ivan Ilyich and Other Stories* (R. Edmonds, Trans.). London: Penguin Books.

Tolstoy, L. (2010). *The Death of Ivan Ilyich and Other Stories* (R. Pevear & L. Volokhonsky, Trans.). London: Vintage Books.

Toolan, M. (1992). The Significations of representing dialect in writing. *Language and Literature, 1*(1), 29-46.

Toury, G. (1995/2012). *Descriptive Translation Studies - and Beyond*. Amsterdam: John Benjamins.

Tymoczko, M. (2012). The Neuroscience of Translation. *Target, 24*(1), 83-102.

Venuti, L. (2008). *The Translator's Invisibility*. Abingdon: Routledge.

Walker, C. (2018). Towards a Quantitative Measurement of Equivalent Effect and a Tentative Conceptualisation of Cognitive Equivalence. In C. Walker & F. M. Federici (Eds.), *Eye-tracking and Multidisciplinary Studies on Translation* (pp. 225–258). Amsterdam: John Benjamins.

Walker, C. (2019). A Cognitive Perspective on Equivalent Effect: Using Eye Tracking to Measure Equivalence in Source Text and Target Text Cognitive Effects on Readers. *Perspectives: Studies in Translatology, 27*(1), 124-143. doi:10.1080/0907676X.2018.1449870

Walker, C. (2020). *An Eye-Tracking Study of Equivalent Effect in Translation: The Reader Experience of Literary Style*. London: Palgrave Macmillan.

Walker, C., & Federici, F. M. (Eds.). (2018). *Eye-tracking and Multidisciplinary Studies on Translation*. Amsterdam: John Benjamins.

Xiao, K., & Muñoz Martín, R. (2020). Cognitive Translation Studies: Models and Methods at the Cutting Edge. *Linguistica Antverpiensia, New Series: Themes in Translation Studies, 19*, 1-24.

Ytre-Arne, B. (2011). 'I Want to Hold It In My Hands': Readers' Experiences of the Phenomenological Differences Between Women's Magazines Online and In Print. *Media, Culture & Society, 33*(3), 467-477.

[1] A useful resource with videos of the majority of these FACS units can be found on the iMotions website at the following address: https://imotions.com/blog/facial-action-coding-system/ [Accessed on 9 July 2020].

[2] 'Volokhonskaya', in the original Russian. I retain the Anglicised 'Volokhonsky' here to maintain parity with the presentation of her name in the Penguin books and in criticism.

[3] For reference, the original Russian and literal back-translations are as follows: «делу Мельвинских» (*lit.* the case of the Melvinskys), «доказывая неподсудность» (*lit.* demonstrating non-jurisdiction), «не вступив сначала в спор» (*lit.* having not initially entered into the argument), and «только что поданные 'Ведомости'» (*lit.* just-delivered Vedomosti [a local newspaper]).

[4] Gaze duration is calculated from the sum of all first-pass fixations within a specific area of interest; total fixation duration is the sum of all fixations within the area of interest (including any regressions and subsequent passes). As a first-pass metric, gaze duration tends to reflect initial text integration processes, while total fixation duration, since it comprises all passes, is indicative of the total cognitive effort exerted over a specific area of interest..

[5] Software such as iMotions can also identify 7 core emotions – joy, anger, fear, disgust, contempt, sadness, and surprise – in addition to valence and intensity, but there is potentially more scope for error in these crude algorithmic qualitative analyses of facial expressions.