Computers and Education: Artificial Intelligence 3 (2022) 100094

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.sciencedirect.com/journal/computers-and-education-artificial-intelligence

Learners' web navigation behaviour beyond learning management systems: A way into addressing procrastination in online learning?

Eduard Pogorskiy *, Jens F. Beckmann

School of Education, Durham University, Leazes Road, Durham, DH1 1TA, United Kingdom

ARTICLE INFO

Keywords:
Web navigation behavior
Online learning
Procrastination
Web browser extension
Sequential pattern mining
Behavior visualization

ABSTRACT

The attractiveness of online games, social media, and mobile apps is frequently considered a challenge for online learners. Procrastinatory behaviour is often associated with a relative lack of self-regulatory skills that would otherwise help learners to resist distractions and to progress in learning. This paper reports a pilot study, conducted with 49 online learners, in which we describe the use of a virtual learning assistant as a tool for collecting online learners' web navigation behaviour. As this virtual learning assistant operates as an extension to the Chrome web browser, it is possible that data collection is achieved independently of, and beyond specific learning management systems. Furthermore, the study opens up the possibility of leveraging the collected dataset for visual learning analytics and pattern mining. To demonstrate the potential utility of the virtual learning assistant, we present an example for a detailed examination of a learner's web navigation behaviour. The results of the detailed examination of a single learner's web navigation behaviour over 333 days, presented as a case study, revealed the presence of seasonality in accessing certain web resources and stable sequential patterns in the learner's web navigation that can be associated with procrastinatory behaviour.

1. Introduction

With the challenges that emerged from the COVID-19 pandemic, many education providers shifted their learning provision to various online environments. This transition exacerbated a known issue associated with online learning environments, namely, increased exposure to distractions (Robal et al., 2018; Svartdal et al., 2020) coupled with a relative lack of support (Reich & Ruipérez-Valiente, 2019). To mitigate the associated risks, learners need to effectively utilise self-regulatory skills — one of the factors crucial to educational success (Reparaz et al., 2020). Self-regulation plays an essential role in navigating the challenges of online learning. It helps learners to master new learning materials, to persist with their study of educational content, and to achieve their ambitions as lifelong learners (Nussbaumer et al., 2015). The research presented here is related to the question of whether success in online learning is linked to individual differences in self-regulation.

Self-regulation is traditionally conceptualised and operationalised as a time and situation stable characteristic (Pogorskiy et al., 2018) that enables learners to utilise their skillset across different conditions and time frames. As noted by Kelley et al. (2015), it requires four mechanisms. The first includes an awareness of one's behaviour in order to be able to compare it to established norms. Second, the learner needs to understand the consequences of their behaviour. Third, the learner needs to be cognisant of the potential repercussions of not performing a certain behaviour. Finally, the learner needs to find a compromise between their own expectations and standards and those that are external, i.e., learned norms (Kelley et al., 2015). While these four mechanisms might work reasonably well in traditional classroom settings, in online settings, especially in the context of those massive open online courses where the emphasis is clearly on instructions with rather limited opportunities for non-instructive social interactions, these four mechanisms, required for self-regulation, might be difficult to implement. As a result, learners may engage in procrastinatory behaviour.

As highlighted by Kelley et al. (2015), experience of emotional and social stress, deficits in self-regulatory resources, and exposure to potentially distracting external cues, such as social media websites which occasionally are characterised as having addictive qualities (Osatuyi & Turel, 2018), may all lead to procrastinatory behaviour. Previous research has demonstrated that procrastinatory social-media usage, particularly the use of facebook.com, is common among learners; it negatively affects students' well-being and increases the chances of experiencing stress in the academic context (Meier et al.,

* Corresponding author.
E-mail address: eduard.pogorskiy@durham.ac.uk (E. Pogorskiy).

https://doi.org/10.1016/j.caeai.2022.100094
Received 6 January 2022; Received in revised form 18 August 2022; Accepted 21 August 2022
Available online 6 September 2022

2016). Facebook usage is negatively associated with the effort learners invest in educational activities (Junco, 2012) and can cause increased levels of anxiety over time (Sternberg et al., 2020). Therefore, it is important to understand the impact of tempting external distractions, such as social media and entertainment websites, on learners' self-regulation and procrastinatory behaviour. This enables not only researchers but also educational practitioners or learners themselves to identify, prevent, and intervene in the potentially negative pathways associated with external distractions found online.

Assessing learners' self-regulation as a state — in contrast to the prominent trait conceptualisation — also provides a more informative way of gaining an understanding of the learning process and may subsequently help to effectively intervene in order to provide support to learners as necessary. One common form of assessing learners' self-regulation and the associated effects of social media use on learners' on-task behaviour and their academic performance is the use of self-reports. Self-reported media use behaviour, such as the use of instant messaging, social media, and micro-blogging, assessed with the Media Multitasking Index-Short (Baumgartner et al., 2017), is associated with lower academic performance based on a survey of $N = 1{,}445$ students from three Southern African countries conducted by le Roux et al. (2021). However, excessive reliance on self-report data has its limitations as research participants might not always be willing to disclose or be aware of their actual behaviour. An alternative approach frequently adopted in research studies is to analyse learners' behaviour using tracking software and physical devices. For example, this includes sensors embedded with wearable technologies (Moissa et al., 2019). Such methods, e.g., self-report and behavioural measures, can be applied simultaneously to cross-validate or to triangulate data collected using other approaches. In one such example, self-reports for assessing procrastinatory behaviour were delivered via electronic diaries using a mobile app as part of an interactive ambulatory assessment (Loeffler et al., 2019). Thus, several approaches can be applied to assess learners' self-regulation and media use, including implementing traditional self-reporting questionnaires, collecting digital behavioural traces, or a combination of both. However, self-reports and behavioural measures obtained solely within a learning management system do not provide a complete picture. There are a variety of interactions between characteristics that are attributed to the learning process that might have an impact beyond what is measured by self-reports and behavioural measures. These include a learner's characteristics, characteristics of a learning task, and a learning situation, as in the three-dimensional framework Person-Task-Situation proposed by Beckmann (2010) (see also Beckmann and Goode (2017)). External distractions to online learning, such as social media website visits, could be attributed to the learning situation in this triad. That is, the situation imposes a stimulus that the person (i.e., learner) engages with at the cost of (cognitive or otherwise) resources not being dedicated to dealing with the actual learning task. Therefore, the focus of this work is on data collection relating to learners' web navigation behaviour in naturalistic or field settings.

With the rise of applicability of Learning Analytics (LA) and Artificial Intelligence in Education (AIEd), collecting behaviour traces is particularly important for predictive analytics due to differences in learners' behaviour, as revealed in an analysis of 3,900 articles concerning LA conducted by Chen et al. (2022). Furthermore, Chen et al. (2020) in their systematic review of 45 highly cited research papers dedicated to utilising AIEd, highlight the need to facilitate the application of deep neural networks in the context of modelling learners' latent processes and their behaviour to build detectors for educational interventions. The authors identify this aspect as one of the current gaps in the research literature. Collecting learners' web navigation behaviour in naturalistic settings can supplement this task by providing an additional valuable data source to fill this gap.

This paper contributes to monitoring learners' web resource usage by describing a practical implementation of a novel data collection tool to capture learners' web navigation behaviour, deliver interventions, and track responses to such interventions beyond the constraints of learning management systems. We argue that certain patterns in learners' web navigation behaviour can be characterised as procrastinatory behaviour. If identified as such, adaptive interventions can be devised and leveraged in order to provide personalised support to students beyond learning management systems. This approach to collecting learners' web navigation behaviour in naturalistic settings could also help gain a better understanding of learners' media use and its role in their on-task and off-task behaviour, thus enabling researchers and educators to develop instruments that can both prevent procrastinatory behaviour and intervene in order to aid learners' self-regulation.

2. Background

2.1. Self-regulation in learning

Self-regulation takes on various forms, allowing for the control of emotions, actions, daily routines, and some mental processes (Ludvigsen et al., 2018). However, this study focuses specifically on the role of self-regulation in learning. In response to the need to specify the processes involved in the self-regulation of learning, the concept self-regulated learning (SRL) has been developed over the past several decades by educational psychologists. There are multiple prominent theories of SRL which focus on learners' achievement, behaviour, and utilisation of strategies to pursue desired learning goals. Influential and established theories include those proposed by Zimmerman (2000), Boekaerts (1999, 2017), Butler and Winne (1995), Winne and Hadwin (1998), and Pintrich and De Groot (1990); Pintrich et al. (2000). These theories have consolidated theoretical and empirical backgrounds and have been broadly acknowledged as established theories of SRL by researchers and educators alike (Panadero, 2017).

Conceptually, learners' self-regulation is not a unitary construct. Rather, it is characterised by different types of self-regulated actions that are dependent on tasks, applied domains, and socio-cultural contexts (Kaplan, 2008). In terms of choosing a specific SRL model as the most appropriate direction for intervention design, Zimmerman's (2000) notion of SRL multidimensionality and the focus on cognitive processes involved in SRL highlighted by Winne and Hadwin (1998) and Butler and Winne (1995) seem the most relevant to the present study.

In contrast to focusing on individual self-regulatory processes, such as goal setting and strategy use, Zimmerman chooses a different approach. His effort to unite distinct elements into a multifaceted construct led to the multidimensional view on learners' self-regulation (Zimmerman, 2008). The multidimensional approach to learners' SRL explains why some learners may self-regulate on a certain task while others experience difficulty. Butler and Winne's model of SRL (Butler & Winne, 1995) is based on theories of information-processing, and this model, as it was initially proposed in 1995, includes four phases: i) external, utilising available resources that are external to the learner, and internal, relying on memory as a resource, information searches relevant to a task at hand; ii) goal setting and the creation of a plan to achieve the set goals; iii) working on the task with the extracted information toward the goal(s); iv) evaluation of progress and goal adjustment, if required (Butler & Winne, 1995). In his more recent work, Winne has identified several basic cognitive processes involved in SRL. These processes correspond to higher level operations performed by learners: searching, providing attention to information, monitoring, identifying suitable information, assembling, combining separate information by identifying relationships, rehearsing, preserving information, and translating, transforming the representation of information provided (Winne, 2017, p. 37).

2.2. Learners' procrastinatory behaviour in online learning

Learners' procrastinatory behaviour is likely to emerge as a result of low levels of self-regulation. Academic procrastination is considered to be a situational and dynamic construct (Ziegler & Opdenakker, 2018) that is associated with a deficit in one or several of the components involved in self-regulation. SRL in online settings in some situations can be utilised as a negative predictor of online learners' procrastination disposition (Cheng & Xie, 2021; Hong et al., 2021), especially time management (Cerezo et al., 2017). A survey of 7400 participants conducted by Steel et al. (2018) to determine the epidemiology of procrastination demonstrated that, in the majority of cases, procrastination could be explained with learners' resources involved in their self-regulation, including attention control, energy regulation, which has been understood to demand significant mental resources, and automaticity (defined as habitualised courses of action that require minimal or no conscious attention). These factors accounted for the majority (74%) of variance in procrastination (Steel et al., 2018, p. 13). Procrastinatory behaviour can be divided into two categories: controlled and uncontrolled procrastination, or as defined by Wessel et al. (2019), active (intentionally delayed) and passive or unintentionally delayed procrastination. Learners might be engaged in controlled procrastination purposefully. For example, in the case of cognitive overload, they might deliberately free the cognitive resources required to accomplish a task by switching their attention to an activity that makes fewer demands on their cognitive resources. Uncontrolled procrastination may occur involuntarily, due to working memory (cognitive) overload, emotional distress, and motivational problems when attempting to engage in certain activities. As such, it can be hypothesised that procrastinatory behaviours can be captured via certain measures and can be tracked as a frequent behaviour that constitutes a pattern.

2.3. Controlled and uncontrolled procrastination

Under the umbrella of "controlled procrastination", it is assumed that, instead of a learning session as such, learners might be engaged in nonetheless beneficial, self-aware procrastinatory activities. This controlled procrastination might occur after a high-intensity or lengthy study session when a learner seeks relaxation, or an activity with low-level cognitive demand — similar to cognitive offloading or reliance on external resources to reduce cognitive demand, as defined by Hu et al. (2019). Some learners can use controlled procrastination as a motivator, e.g., after studying for 1 h, learners might allow themselves 10 min of "Facebook time". Used in this way, these activities can be considered productive and useful tools for learning.

"Uncontrolled procrastination" occurs when learners are unintentionally engaged with counterproductive activities. Such procrastinatory behaviour could be attributed to several causes, for example, experiencing high levels of stress. Research suggests that learners increasingly experience mental health problems, with anxiety and depression prevalent among graduate students (Evans et al., 2018), as well as those in primary, secondary and further education (Tremblay et al., 2011). A significant number of school-age children have been found to have low self-esteem, alongside problems associated with excessive sedentary behaviour, screen-time, and extensive use of social media (Tremblay et al., 2011). Time spent on social media and overall screen-based media interactions significantly correlate with a decline in well-being among young people, which appears to have an effect on their long-term performance at school and life outcomes. In particular, this seems to be the case for female pupils (Booker et al., 2018). A systematic review of published studies (Suchert et al., 2015) and a meta-analysis of observational studies (Liu et al., 2016) show that screen time and screen-based sedentary behaviours are connected to anxiety and depressive symptoms, inattention, problems with hyperactivity, low self-esteem, a low sense of well-being and overall quality of life. Although little is known about the proportion of online learners who experience symptoms related to anxiety and depression, it can be estimated that the nature of online learning environments — with the absence of university health services, reduced instructor and peer support, prevalence of exposure to screen time and sedentary behaviour — anxiety and depression are likely to be at least as typical as for school-age children as it is for students enrolled in graduate-level courses. This assumption can be traced to emerging research on the topic, for example, recently published protocols of randomised controlled trials aimed at evaluating the effectiveness of internet and app-based stress interventions for distance-learning students with depressive symptoms (Harrer et al., 2019) and an internet-based intervention to address procrastination in college students (Küchler et al., 2019).

Based on the assumption that a significant proportion of online learners may experience problems with their self-regulation in online learning environments due to anxiety, it is crucial to understand how this issue may affect the learning process, and what effects it might have on learners' engagement with educational resources. Based on research in psychology and neuroscience, a dynamic framework for understanding mind-wandering has been proposed by Christoff et al. (2016). This framework links mind-wandering to depression and anxiety, characterised by one's involvement in repetitive, automatic actions (Christoff et al., 2016, p. 725).

We, as have others, argue that procrastination as consequence of problems with self-regulation manifests itself via repetitive behavioural patterns. Therefore, it might be worth attempting to track repetitive patterns as part of the process of identifying learners' involvement in uncontrolled procrastination, that negatively affect learners' engagement with their online learning environment. Some learners may experience problems dealing with the affective, cognitive, metacognitive, and motivational demands of online learning and may develop symptoms related to depression and anxiety. In these cases, uncontrolled procrastination is considered a counterproductive behaviour.

2.4. Approach to collecting procrastinatory behaviour data

Behavioural traces captured whilst interacting with an online learning environment and web browsing in general can be informative of a learner's level of self-regulation. Learners perform actions in their web browsers: they open tabs in their browser windows, visit URLs, switch between opened tabs, and switch between their browser and other installed software. Each of these actions can be considered as a single point of activity. For example, a learner might open an online course website on the online learning platform "edx.org" in their browser, spend 1 min on this URL without interruption, and might then open another website, e.g., "hollis.harvard.edu", spending another minute on this second page. This sequential activity consists of two events. In this example, both activities can be considered learning-focussed.

With the obvious exception of traces drawn from single events, it is essential to characterise traces as sequences of events. Sequential events are taken together to form time-series data. Such time series data can provide the basis for gaining insights into learners' engagement in learning and in their interactions with the given learning environment, both of which are assumed to depend on processes of self-regulation. Web navigation happens across browser windows and in single window tabs. Some learners may use several different browsers concurrently, alongside additional software installed on their machines. However, using two or more different browsers concurrently is assumed to be relatively uncommon, while software use can be characterised as being less disruptive in comparison to the web environment. It should also be acknowledged that online learners will not necessarily spend all of their time in front of their laptop or other electronic devices and might be distracted by other external events whilst their browsers are open. Learners could even leave their electronic devices with opened web pages in idle mode. Traces captured during the mentioned scenarios should be considered noise and need to be analysed or interpreted with caution.

To answer our main research question of whether success in online learning is linked to individual differences in self-regulation, we aim to address the following sub-questions:

- How are short-term and long-term behaviour patterns manifested in recorded web-browsing sessions?
- How can web-browsing behavioural patterns of an individual be linked to procrastinatory behaviour?

3. Methods and materials

3.1. Participants

The data used in this study were collected during the period from 15 January 2018 to 9 January 2019 as a pilot for a larger scale study. The dataset includes information about voluntarily participating online learners who were enrolled in a diverse range of massive open online courses. The sample ($N = 49$) includes participants with an average self-reported age of 32 years ($N = 14$, $M = 31.71$, $SD = 7.29$). Participants indicated that they completed at least undergraduate ($N = 8$), post-graduate ($N = 4$), and doctoral degrees ($N = 1$). One participant indicated their completion of at least secondary or high school, and one participant mentioned “other education” as their highest level of education achieved. 15 participants self-reported their gender, with the majority being male ($N = 11$). 14 participants indicated their country of origin, with four participants from the United States. Other countries of origin include the Russian Federation, Turkey, Nigeria, the United Kingdom, Australia, Mexico, Spain, Ukraine, India, and Chile. However, not all participants revealed their demographic information. The demographic questionnaire was voluntary, and the remaining participants skipped some questions, opting not to disclose their demographic information.

3.2. Materials

To collect data on learners’ web navigation behaviour beyond the reach of a particular learning management system, the concept of the virtual learning assistant — implemented as an extension to the Chrome web browser — was utilised (Pogorskiy et al., 2018). The practical implementation of this tool consists of three components, (1) a web application with a user interface that enables goal setting, progress monitoring and self-evaluation, (2) an extension to the Chrome web browser to collect trace data and display notifications (Fig. 1), and (3) a Structured Query Language database as a repository for collected trace data.

The web application of the assistant includes several sub-components that allow learners to interact with the tool. First, there is a goal setting interface where learners can indicate one or several goals in terms of an online course or courses they wish to complete. Learners can both indicate a start date and set a deadline for their goal, indicating the time range required for the completion of a given course. Learners are encouraged to provide information on any course-related discussion forums, if there is one linked to the course, the proportion of the course which has been completed to date, that can be adjusted at a later stage, the intended time commitment towards the goal, and to indicate the course name, which will appear in their list of added courses (learning goals).

Another sub-component of the web application fulfils self-monitoring and self-evaluation functions. This sub-component presents a dashboard to support the self-monitoring function to learners, alongside feedback statistics indicating their behaviour recorded by the web browser extension. The summary statistics of time spent by a learner on each online web domain is calculated each day, with daily and weekly statistics displayed in real time to learners, providing feedback on their behaviour. The self-evaluation functionality of the tool is also supported

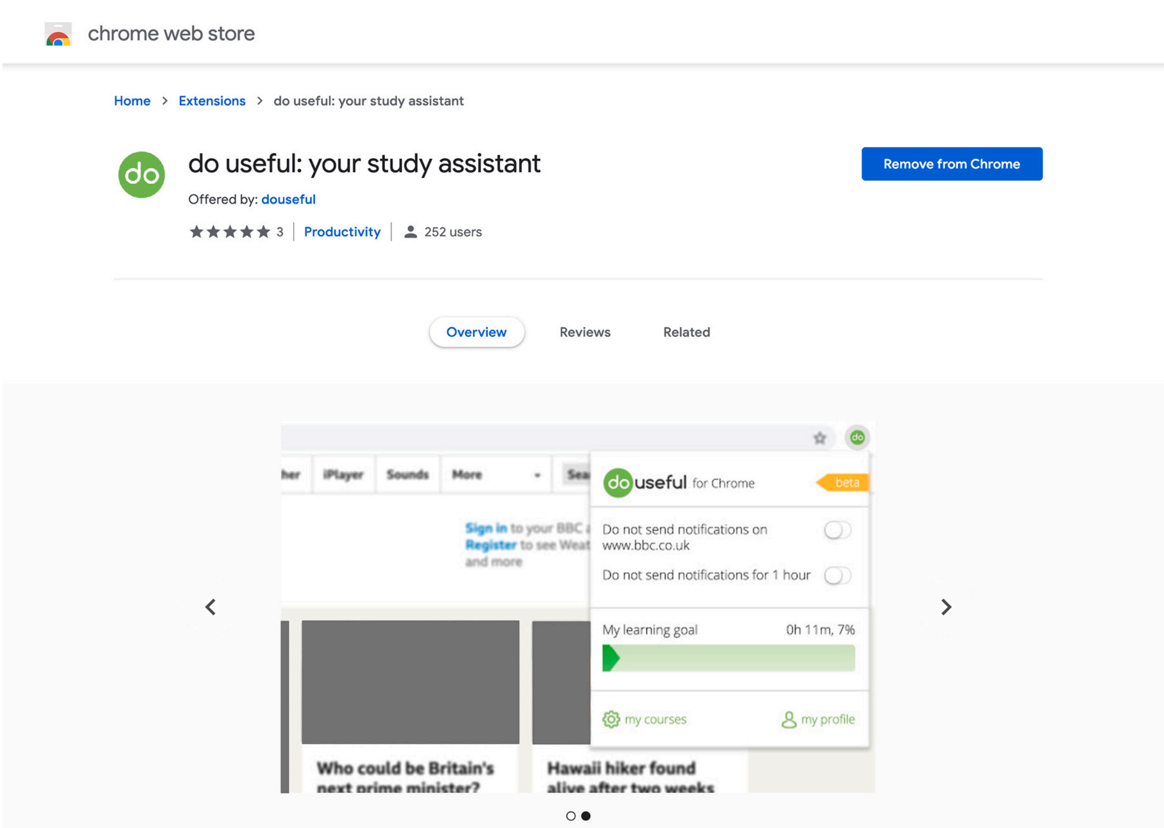

Fig. 1. Virtual learning assistant offered to users of Chrome web browser.

with a dashboard. This dashboard provides learners with a visualisation of summary statistics showing the time committed by a learner to their indicated course (their learning goal). The desired time is indicated next to recorded and displayed summary statistics, allowing the learner to evaluate their progress towards their goal, which was entered during the goal-setting stage.

The virtual assistant also includes an intervention component in the form of pop-up messages that appear in learners' web browser environments in response to learners' behaviour. This component of the tool is based on several decision rules that determine the appearance of distinct templates with text content on the learner's screen. The provision of onscreen messages within the web browser environment is illustrated in Fig. 2. The onscreen pop-up messages are personalised based on observations specified in this figure ("Observations") and the decision rules described in the figure's middle section ("Available?").

It is assumed that behaviours that are likely to represent issues with self-regulation are expressed by extensive time spent on resources that are not related to the indicated learning goals. These moments are calculated based on time spent on web domains captured within the browser extension. When such procrastinatory behaviour occurs, an intervention message is displayed to encourage a shift in learners' behaviour. To achieve this, behaviour tracking functionality was implemented in the tool. Fig. 3 provides a schematic summary of the tracking components of the virtual assistant and a hypothetical scenario of a learner's behaviour, alongside actions taken by the adaptive assistance component of the virtual assistant triggered by the learner's behaviour.

As can be seen from Fig. 3, learners may dismiss or accept a displayed textual prompt shown on their screen, indicated by arrows with the corresponding text "negative response" and "positive response". In the case of acceptances, the learner is redirected to their indicated online course web page in a new web browser window. Learners' web navigation behaviour and responses to each displayed prompt are recorded in the tool's database. Table 1 provides an example of behaviour trace records relating to learners' web navigation behaviour.

3.3. Procedure

3.3.1. Data collection

In this study, participants were invited to use the virtual assistant and to participate in the study in several ways, including promoting the tool among online learners in three Facebook groups, and displaying the information about the data collection tool in the Chrome web store, a catalogue of extensions to the Chrome web browser, as illustrated in Fig. 1. In collaboration with the Higher School of Economics, a link with information about the tool was sent to learners enrolled on the course "Communication theory: bridging academia and practice" offered on coursera.org.

After installing the extension and creating an account on the virtual assistant website, learners were asked to set their goals by providing details of the online course or courses that they wished to complete, and the amount of time they wished to spend per week in order to achieve their goal. Participants were also given the opportunity to personalise their experience by adjusting a list of websites where notifications appeared more frequently, websites where notifications were not shown, and a list of "incognito" websites that were not to be tracked.

Data collected from learners who installed an extension to their web browser and created an account on the project website was found to vary, from a few hours of using the data collection tool to 335 days of recording web session activities, and 35 participants used the assistant for more than 7 days. The collected data include information regarding learners' demographic characteristics, as learners were asked to indicate their country of origin, age, gender, and their level of received education, details about courses they wish to complete, the amount of time that they desired to study per week on the course web page, information about web sessions expressed in time spent on different web domains, and sequences of visited websites.

The final dataset consists of information relating to 361,358 records relating to learners' web sessions, including information on anonymised user IDs, visited domain names, e.g., "ed.ted.com", without specifying the full visited URL addresses to ensure the participant's privacy, adhering to the ethical standards of the study, and a timestamp. Further, this dataset includes 3,649 records of the participants' responses to 11 different on-screen notification templates. These responses to notifications include information about approval or rejection status.

3.3.2. Data labelling and pre-processing

To prepare the raw data for analysis, several steps were taken to shape the data into a suitable format, and to allocate additional data features. First, the list of participants' unique visited websites was sorted according to the frequency of the URLs' appearance in the dataset. Second, in addition to the URLs indicated by participants as their course websites, and categorised as educational URLs, 350 unique frequently appearing website domain names, extracted from a total of 13,284

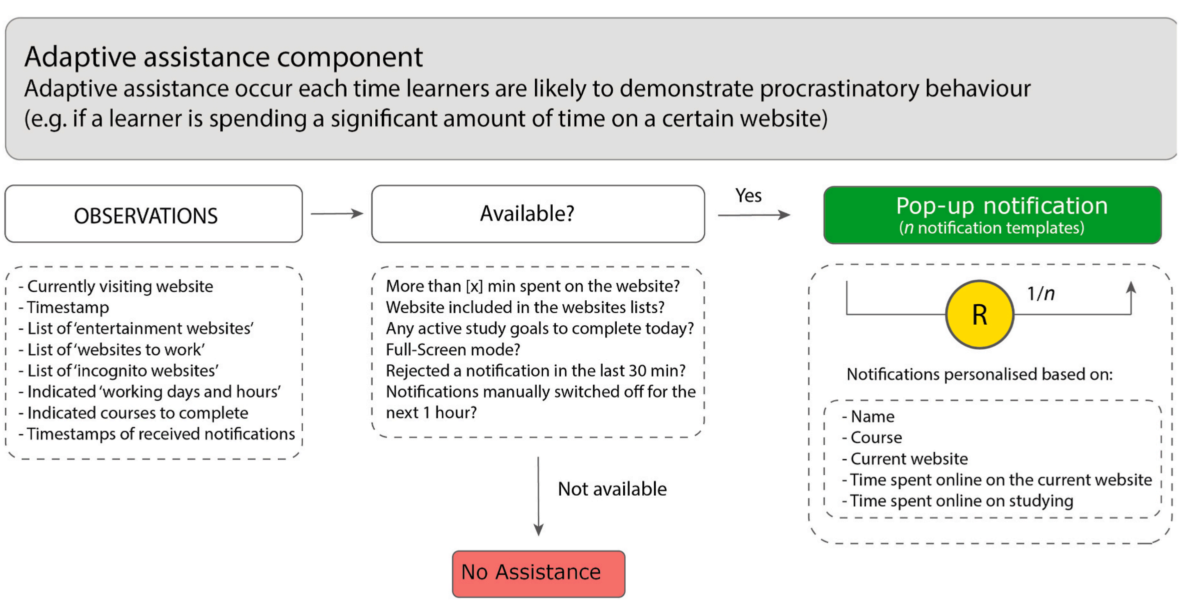

Fig. 2. Schematic illustration of the intervention component and its decision rules implemented within the tool.

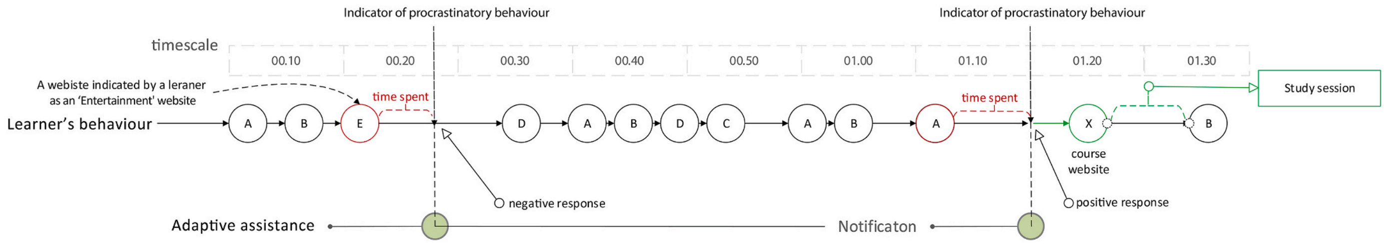

Fig. 3. Schematic illustration of a hypothetical learner's behaviour and an example of the compensatory functionality of the tool as the response to observed behaviour and the occurrence of the procrastinatory behaviour.

Table 1
Example of behaviour traces to illustrate the data structure of the collected dataset.

User ID	URL	Seconds	Timestamp
…			
21567	edx.org	102	2018-07-21 18:50:48
21567	pythontutor.com	2	2018-07-21 18:52:30
21567	facebook.com	4	2018-07-21 18:52:32
21567	cnn.com	10	2018-07-21 18:52:36
21567	netflix.com	196	2018-07-21 18:52:46
…			

unique URLs, were manually coded according to the website categories they might represent. In total, 450 domain names, including course URLs, were labelled according to six distinct categories: "Education", "Media", "Productivity", "Search", "Shopping", and "Social media" websites. Although it only covered a small proportion of the full set of websites, this subset amounted to 46% of all records in relative terms, or 70% of participants' total time spent online, after taking into account the duration of recorded time associated with each URL.

URLs categorised as "Education" correspond to either websites denoting a course to complete, or sites that can be directly attributed to online learning, learning materials, educational content, or university websites. Frequently listed URLs ending with ".edu" or ".ac.uk" were, for instance, placed in this category. Websites categorised as "Media" include domains such as youtube.com, netfilx.org, and websites associated with online gaming. Websites categorised as "Productivity" include domain names that can be indirectly linked to online learning, such as online libraries, translating services or online document editing websites. Websites categorised as "Search" include common search engine websites, such as google.co.uk. Websites categorised as "Shopping" include websites linked to online retail, such as amazon.com, ebay.com, and walmart.com. Websites categorised as "Social media" include domains such as facebook.com and instagram.com. Websites that were not included in the above-mentioned categories were assigned to a seventh category: "Other" websites.

To examine the participant's behaviour for the presence of any patterns that could lead to the abandonment of the course, we selected URL pairs with the latest record of the course URL visit (e.g., course URL — udemy.com) and a following this visit URL ("exit" URL). Thus, we identified URLs that appeared after visiting the course URL where the participant did not return to their course website on the same day, and only for days with a visit to the course URL. Therefore, we selected subsequent "exit" URLs linked to the latest learning sessions, after which the course URL did not appear in the participant's web navigation behaviour for each day with the indicated online course domain visits.

4. Results

4.1. Long and short-term general behavioural patterns

The online tool presented in this study allowed for the extraction of summary statistics on a group level aggregate, but also on a more in-depth examination of individual participants' behaviour. The results of the study described in this section includes a general overview with summary statistics for all participants, presented in Table 2, alongside a case study with a detailed exploration of one individual's behaviour. We first provide an example with summary statistics that include the number of courses among all participants, days participated in the study, and unique URLs visited by participants (Table 2). The summary statistics can be used to identify general patterns across participants, for example, on average, participants visited 914 unique URLs, as can be noted from the table.

4.2. Detailed examination of individual behavioural patterns

In this subsection, we present an example for an in-depth examination of behaviour patterns for a single learner. This case study perspective includes visualisations of the participant's time allocation for different categories of web resources and identified web-navigation behaviour patterns. It also provides a detailed illustration of the selected learner's behaviour in their web environments over a period of 333 days is provided as an exemplar for the possibilities the online tool provides for the analysis of SRL behaviour. For that purpose, we selected a participant whose behavioural observations were collected over a prolonged period of time. Consequently, all participants were ranged according to the duration of their enrolment in the study. Two participants who used the assistant for a prolonged period of time (roughly one year) were selected, each respectively with 335 and 333 days in total. A brief summary check revealed that the first learner who participated in the study throughout the 335-day period added a total of 54 courses, 35 of which were short online courses provided on linkedin.com. This behaviour was considered to be unusual, given that participants rarely indicated more than two courses for completion, as presented in Table 2. Therefore, the participant with the second longest period of use, 333 days in total duration and 298 days with records, with three courses added to their list, was selected for the case study.

Web browsing behaviour can be characterised by the prevalence of certain websites in a learner's web browsing history. For this participant, the ten most frequently accessed URLs, calculated according to days with appearances, where there were 298 days with records, include: google.com (270), facebook.com (251), mail.google.com (206), youtube.com (188), udemy.com (139), amazon.com (125), github.com (121), stackoverflow.com (117), guru99.com (114), and nypl.org (101). It can be noted that social media, e.g., facebook.com, and video hosting, e.g., youtube.com, websites constituted the learner's daily routine, along with the use of a search engine and email services provided by google.com. The participant's time commitment to the selected frequently appearing URLs is illustrated in Fig. 4.

Web browsing behaviour presents seasonality in accessing certain web resources and regular daily routines. Hourly, daily, weekly, and monthly trends presented in Figs. 5–7 provide a visual representation of traces collected within the browser extension. Daily time commitments to the labelled website categories for the selected participant are visualised in Fig. 5. In this figure, we get a visual representation of time spent online, measured across hours, days, and months, revealing short-

Table 2
Descriptive statistics of collected behavioural traces for all study participants.

	N	Mean	*SD*	Min	25%	50%	75%	Max
Days in the study	49	44.16	77.25	0.00	5.00	15.00	37.00	335.00
Courses added	30	4.00	9.94	1.00	1.00	1.50	2.00	54.00
Desired weekly learning time (hours)	30	23.84	39.55	1.00	2.88	9.00	24.25	182.70
Productivity URLs added	32	5.81	15.50	1.00	1.00	2.00	3.25	88.00
Entertainment URLs added	15	2.27	1.28	1.00	1.00	2.00	3.00	5.00
Unique URLs visited	46	913.72	1842.30	2.00	63.50	291.50	726.50	8858.00
Time spent per course (minutes)	24	248.90	453.27	0.08	3.91	18.73	225.41	1714.63
Records with learning sessions	24	96.21	180.61	1.00	4.00	15.50	89.75	749.00
Duration of a learning session (minutes)	24	2.25	2.33	0.00	1.00	2.00	3.25	11.00
Time spent on discussion forum	3	108.36	185.08	0.78	1.50	2.22	162.14	322.07
Records with discussion forum sessions	3	120.33	193.85	1.00	8.50	16.00	180.00	344.00
Duration of a forum session (minutes)	3	0.67	0.58	0.00	0.50	1.00	1.00	1.00
Received notifications	26	95.96	154.83	1.00	5.50	27.50	99.75	498.00
Accepted notifications	20	4.50	4.84	1.00	2.00	3.00	4.25	21.00
Accepted notifications (%)	20	17.26	22.43	0.60	3.30	5.10	23.20	75.00

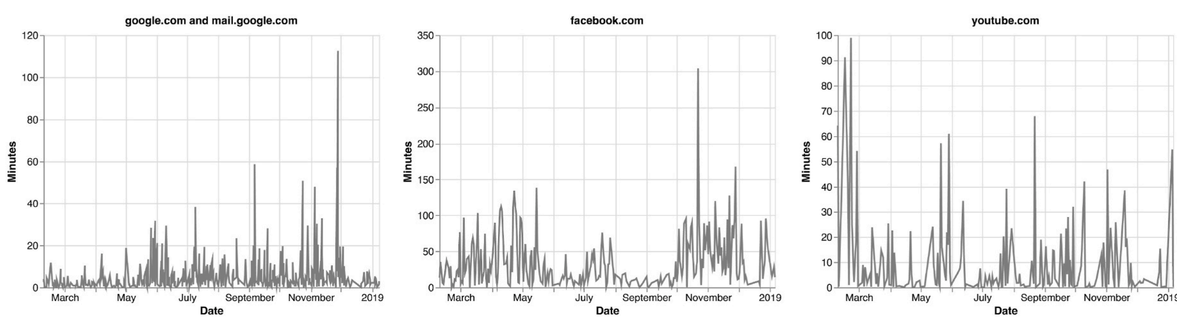

Fig. 4. Time spent by the participant on the selected frequently appeared websites.

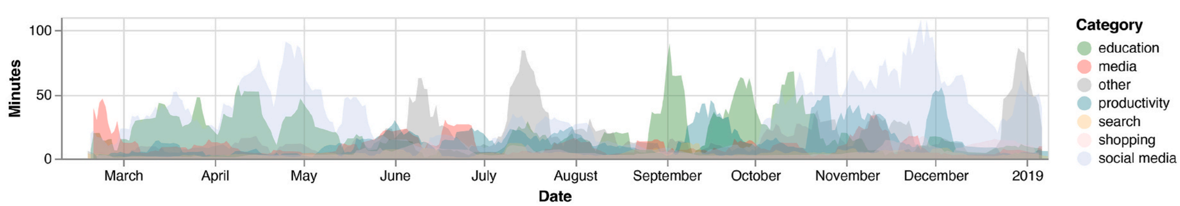

Fig. 5. Example of data visualisation for the participant's time allocation during web navigation behaviour (with the calculated 7-day simple moving average).

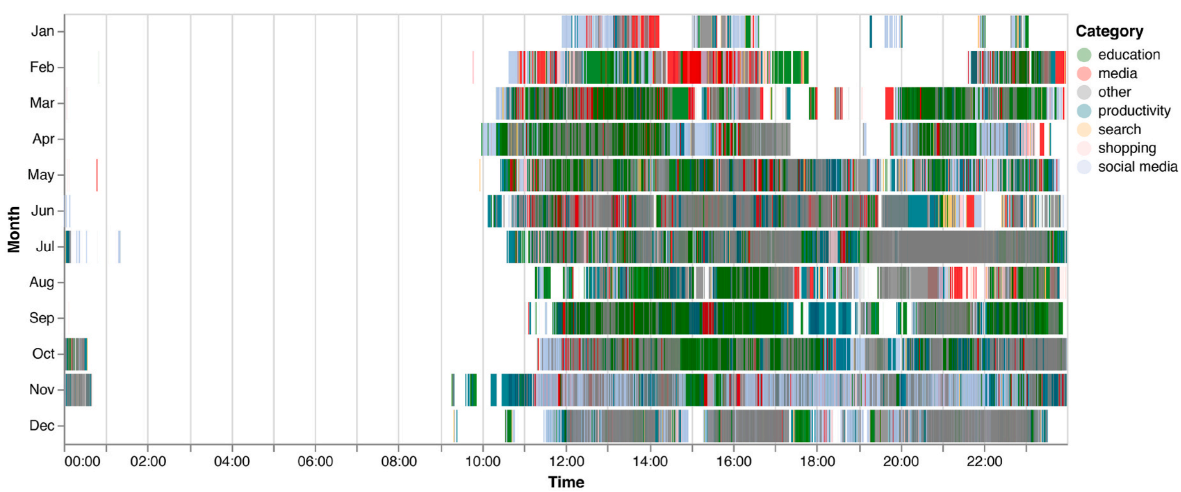

Fig. 6. Example of data visualisation for the participant's time allocation during web navigation behaviour.

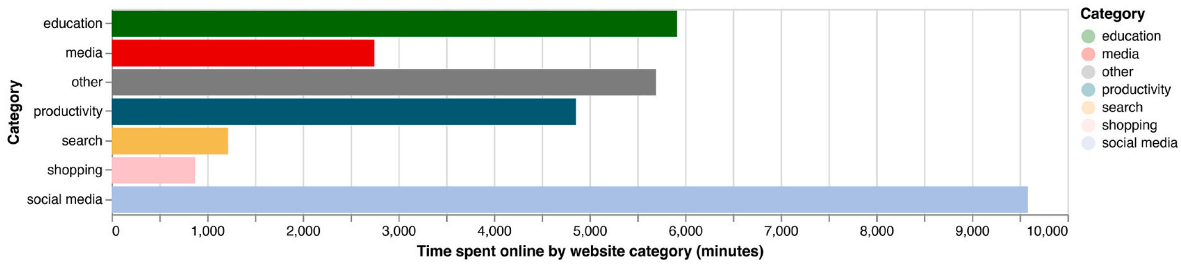

Fig. 7. Example of data visualisation for the participant's time allocation during web navigation behaviour.

term daily trends and seasonal patterns. The figure provides the calculated simple moving average for a 7-day period of time spent online on the selected URL categories. For visualising the data and improving its readability, calculating the simple moving average was helpful in overcoming the burden of extreme values recorded, e.g., in the case of quick swaps between opened web browser tabs or daily prolonged web sessions. This figure allows for trends in web navigation behaviour over the year-long period to be revealed. For example, a seasonality in the participant's time commitment to social media websites can be seen here.

Fig. 6 presents the participant's daily web navigation behaviour routine. There were barely any web browsing activities during the night-time hours for obvious reasons. Furthermore, typical web-access hours can be estimated from this graph, such as the beginning of web navigation behaviour at around 11 a.m. for each day with records. In addition, a handful of unusual web browsing events can be spotted, for example, access to websites categorised as media after midnight in May.

Fig. 7 represents the total time spent online on web domains grouped according to their categories. The neglected role of online retail and media websites relative to the dominant role of social media websites in the participant's web navigation can be observed here.

4.3. Behavioural patterns and procrastinatory behaviour

In this subsection, we show that procrastinatory behaviour can be addressed in real time via pop-up messages. Data collected regarding the participant's responses to the provided interventions can supplement the visualisations presented in Figs. 5–7. In addition to providing an overview of the participant's short-term behaviour as the response to intervention (approving or dismissing), the behaviour tracking capabilities of the tool and visualisations of the collected data provide insights relating to the extended effect of the intervention on the participant's web navigation behaviour. Figs. 8 and 9 are an illustration of merging the participant's web navigation behaviour with their responses to the tool's intervention, with the intention to support the participant's self-regulatory behaviour. For example, Fig. 9 illustrates that a negative response to a prompt was not necessarily associated with the absence of any effect. As can be noted from the figure, despite the participant's dismissal of the notification, their web navigation behaviour stopped, indicating that the participant had at least switched their attention beyond the scope of the web browser, or left their personal computer entirely for at least a period of 30 min following the intervention.

Time dedicated to online courses differs during the year for the participant. As it can be seen from Fig. 5, the distribution of time dedicated to educational web resources differs during the year with periods of peaks and troughs. As mentioned earlier, the selected participant indicated three online courses they wished to complete. The summations of the total time recorded that can be attributed to time spent on the indicated course URLs reveals that the participant did not commit their time resources to all these three courses equally. The recorded traces show that the learner spent, in total, 23.55 h on udemy.com, 16.68 h on freecodecamp.com, and only 16.14 min on their online course on edx.org. The participant's time commitment to the indicated online course URLs is presented in Fig. 10. In order not to increase the complexity of the subsequent analysis, we selected the participant's web navigation behaviour relating to the online course hosted on the udemy.com online learning platform for further exploration, since it was the longest recorded behaviour associated with this course. This approach also enables the study of days with learning sessions, indicated by course URL visits, and the extraction of URL subsets per learning session for each day, as well as the frequency of domain names that appeared after learners ended their learning sessions.

It was noted that the participant frequently revisited the same websites at the end of each online learning session. In Table 3, the various URL pairs associated with the last course URL visit suggests a recurring theme. This table presents "exit" URLs and their frequency of appearance (days). What is striking about this data is that of the total 139 days when the udemy.com course URL was accessed, the participant did not return to the course website after visiting facebook.com on 42

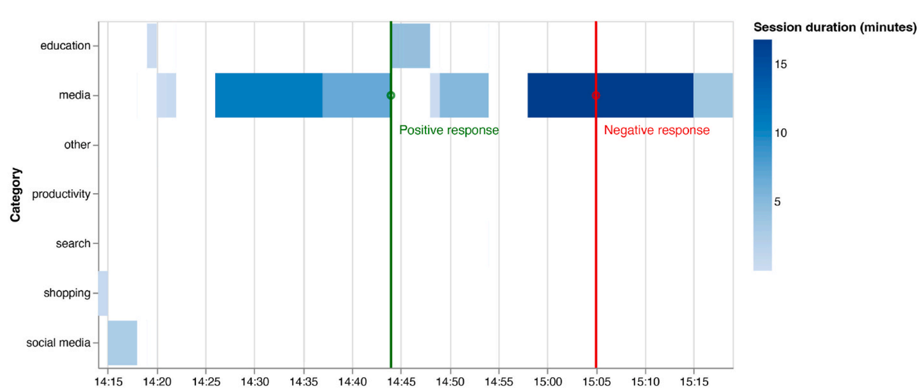

Fig. 8. Example of on-screen notifications, where green for accepted and red for dismissed notifications, and changes in the learner's web navigation behaviour (switches between web domain categories following notifications). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

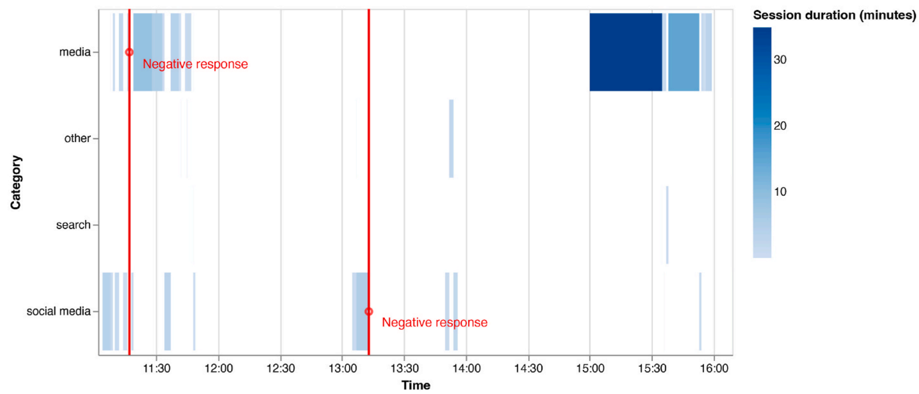

Fig. 9. Example of on-screen notifications and changes in the learner's web navigation behaviour (switches between web domain categories following notifications on different categories of web domains).

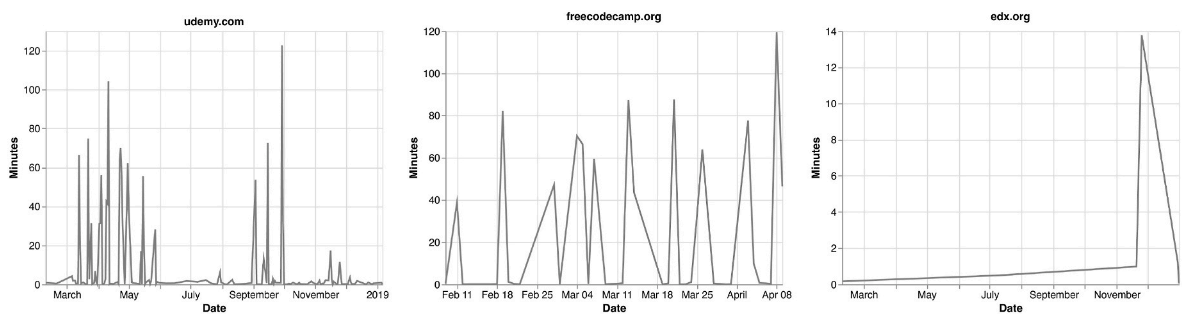

Fig. 10. Time spent on online courses added by the participant to their profile as desired to be completed.

Table 3
Frequency of URLs visits following the latest session for the course website for each day with study sessions.

Course URL	"Exit" URL	Frequency
udemy.com	facebook.com	42
	youtube.com	12
	guru99.com	9
	google.com	7
	beta.freecodecamp.org	3
	medium.com	3
	blogpost.com	2
	dzone.com	2
	hackernoon.com	2
	mail.google.com	2
	oracle-dba-online.com	2
	seleniumeasy.com	2
	slack.com	2

days and after visiting youtube.com on 12 days.

The participant frequently visited a sequence of URLs at the end of learning sessions. In the final step of the analysis, we extended the sequence of URLs following the latest course session for each day, incorporating visits to the course URL and 20 subsequent URL records. We utilised the CM-SPAM (Fournier-Viger et al., 2014), a sequential pattern mining algorithm based on the SPAM algorithm (Ayres et al., 2002), accessed through the SPMF open-source data mining library (Fournier-Viger et al., 2016) to the resulting subset to identify frequently appearing sequences of repetitive behaviour that follow after the learning sessions have ended. The results of applying the CM-SPAM algorithm to examine the derived subset for the presence of any sequential patterns are summarised in Table 4. Patterns that appeared with a frequency of less than 5% of the days in consideration, with a length of fewer than 3 items were excluded from the results. Thus, the results presented in Table 4 demonstrate the relatively stable web navigation patterns that followed the participant's last course URL visit within a learning session for each day. For example, the sequence of 20 URLs that include youtube.com, facebook.com, and google.com followed the latest session on the indicated course udemy.com on 18 days out of the total 139 days with the course records.

Table 4
Sequential patterns in the learner's web navigation behaviour.

Course URL	Item 1	Item 2	Item 3	Support
udemy.com	youtube.com	facebook.com	google.com	18
	facebook.com	google.com	mail.google.com	11
	nypl.org	catalog.nypl.org	browse.nypl.org	7
	facebook.com	google.com	guru99.com	7

5. Discussion

In this study, we utilised a virtual learning assistant, operated through the Chrome web browser to collect data on learners' web navigation behaviour in naturalistic (field) settings. Instruments of visual LA are commonly applied when providing feedback to learners on their behaviour in order to boost meta-cognitive, motivational, affective, and the cognitive aspects of learners' self-regulatory skills (Vieira et al., 2018), as well as expanding an understanding of the processes underpinning students' regulation of their learning (Noroozi et al., 2019). The application of visual LA to collected and synthesised data allows short-term and long-term patterns to be revealed in the case study, for example, the seasonality of social media use and daily web navigation routines. The categorisation of URLs into seven general groups, as applied in this study, helped to differentiate between behaviours that can be attributed as on-task versus off-task activities, it also enabled repetitive behaviour that may signal procrastinatory behaviour to be visualised.

Theoretical implications that emerge from leveraging collected trace data to analyse learners' behaviour are linked to identifying the most frequently appearing web navigation behaviours. A common avenue of research in online education seeks to identify learners' navigational patterns within a single course or on a particular online learning platform (Bannert et al., 2015; Eynon et al., 2016; Maldonado-Mahauad et al., 2018; Rizvi et al., 2020). However, pattern mining beyond the scope of the data provided within learning management systems and in the context of a wider environment — learners' general web browsing navigation — distinguishes this work from previously published research. In this study, we observed sequential navigational patterns that can be interpreted as learners' procrastinatory behaviour, most frequently expressed by use of social media and video hosting websites. The patterns revealed in this case study resonate with the view of procrastination as a situational and dynamic construct and emphasise the relevance of interventions to address learners' procrastination (Ziegler & Opdenakker, 2018).

These findings supplement previous studies that have considered the potentially negative effects of procrastinatory behaviour on academic performance due to extensive use of smartphones and other electronic devices (Breems & Basden, 2014; Inie & Lungu, 2021; Li et al., 2020). The results of this pilot study help to further shed light on the complexity of interactions between the general characteristics of the learning environment, specific interventions provided, and learners' responses to these interventions in the given environment. We employed analysis techniques commonly used in the fields of LA and AIEd (Chen et al., 2020, 2022), such as visualisation and pattern mining, to the collected trace data. Conceptually, we linked these patterns to procrastination. The case study of a single learner's web navigation behaviour demonstrates the usefulness of extensions to web browsers in collecting trace data. It also illustrates how visualisations can be utilised in analyses of behaviour in online learning environments beyond the scope of learning management systems. The information gathered and synthesised by this tool can also inform the intervention strategies necessary to support online learners with self-regulation whilst dealing with online learning opportunities. However, we want to highlight that the purpose of this study was a descriptive proof of concept, rather than claiming any substantive generalisability of the findings. In addition, the issue of problematic procrastination goes beyond online learning, as procrastination has also been identified as a key challenge to the online component of blended learning, among technological literacy and technological accessibility (Rasheed et al., 2020).

The practical utilisation of the learning assistant to trace learners' web navigation data and responses to interventions reported in this study extends previous efforts to utilise social media data, analytics, and visualisation tools available to researchers, such as the Social Media Macroscope (Yun et al., 2020), used for the in-depth examination of data relating to social media, or Clickstream (Filvà et al., 2019), a tool designed to analyse the flow of clicks on a website. In terms of pedagogical implications, data collected with the application of extensions to web browsers can be utilised to support an instructor's role in identifying, selecting, and recommending additional learning resources for learners. Longitudinal learning data provided to teachers and course instructors has a long tradition of being associated with providing benefits for learners at risk (Herodotou et al., 2019). For instance, for the exemplar learner presented above, embedding reading materials into the course platform may have helped to sustain engagement with the course materials. It is assumed the learner was accessing reading materials on nypl.org, as indicated in Table 4. This would have meant less need to access external resources, where additional distractions could have jeopardised the ongoing engagement in the learning session.

Two main limitations can be identified regarding the study presented. First, we did not provide an in-depth examination of web navigation behaviour patterns regarding the remaining study participants as additional case studies and focused solely on one learner with one of the richest data sources available from the enrolled participants. Admittedly, this approach limits the potential for "far-reaching" generalisations, which we did not intend to offer. The aim of the study rather was to provide a proof of concept.

The second limitation of the study could be seen in terms of data accuracy. Data collection in naturalistic settings implies a number of risks for data quality that tend to lie outside the control of researchers (Arechar et al., 2018). For example, with the application of the extension to participants' web browsers, there is a risk that several members of one household may have used the computer with the extension installed. The potential solution to mitigate this issue in future studies is to include a screening question to determine if any other person uses the computer on a regular basis. However, this will naturally limit the variability of potential participants, such as households with limited access to personal electronic devices and may result in the problem of ecological validity of received findings by excluding groups of learners from a certain socioeconomic background. This trade-off should be considered in future studies.

In future studies, any identified repetitive behaviours, based on individual learners' web navigation patterns, can be supplemented with a questionnaire provided to learners that is able to monitor their latent states, such as motivational and emotional aspects of self-regulation. Learners' self-reports can then be linked to behavioural acts that may follow specific sequential patterns, such as visits to URLs. Learners' self-regulatory behaviour actions can be identified by analysing learners' behaviour at a fine-grained level, as described by Greene and Azevedo (2009) based on the work of Azevedo et al. (2008). These actions include performed events that can be linked to learners' self-regulation: planning, e.g., setting goals by making a list of online courses, monitoring, e. g., monitoring one's progress towards a goal, strategy use, e.g., selecting a new source of information, and task difficulty and demands, e.g., help-seeking behaviour (Greene & Azevedo, 2009, pp. 25–27). The analysis of behaviour traces (Siadaty et al., 2016), or self-report measures using questionnaires, or even post hoc interviews, can help to gain a better understanding of processes leading to and managing procrastination in learning contexts. For instance, Cleary et al. (2012) discusses the possibility of accessing SRL at the micro-level using interview data (SRL microanalysis technique). The application of microanalysis to assess individuals' self-regulatory processes can be traced back to Bandura's microanalysis (Bandura et al., 1980), which was used to evaluate shifts in self-efficacy beliefs and the relationship between these shifts and behaviour performance in response to anxiety-reduction interventions (Cleary & Callan, 2018, p. 340). A better understanding of behavioural patterns preceding procrastinatory behaviour might also help to supplement intervention with prevention. In addition, in future studies, other analytical methods and instruments could be utilised for furthering learning behaviour analyses. For example, the SuperNoder tool (Dessì et al., 2018) could be useful in creating an overarching aggregation of identified patterns and nodes to represent modular structures in networks of visited web domains.

6. Conclusion

Online learning happens in online environments: learners' interactions with their environments result in digital footprints. These footprints (i.e., traces) include single events, sequences of activities, and patterns. Learners apply a broad range of actions prior to, during, and after engaging in learning processes, and it is possible to trace such actions. This work contributes to the topic of automated intelligent support in online education applications through its demonstration of the practical use of a tool for capturing learners' behaviour traces and providing interventions beyond learning management systems. This study demonstrates that collected data containing learners' web navigation behaviour traces can be leveraged in order to identify patterns that can be associated with procrastinatory behaviour by applying various visualisation and pattern mining techniques. Insights into how and when learners interact with different online resources, can be a useful source of information for providers of online learning, but also a useful source of feedback for learners to regulate their utilisation of online learning opportunities more effectively.

Ethics

Prior to data collection, each participant was asked to read an information sheet outlining the study. They were asked to agree with a written declaration of informed consent. Ethical approval for this study was obtained from Durham University's School of Education (date of approval: 6 December 2017).

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Arechar, A. A., Gächter, S., & Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics, 21*, 99–131. https://doi.org/10.1007/s10683-017-9527-2

Ayres, J., Flannick, J., Gehrke, J., & Yiu, T. (2002). Sequential pattern mining using a bitmap representation. In *Proceedings of the eighth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 429–435). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/775047.775109.

Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research & Development, 56*, 45–72. https://doi.org/10.1007/s11423-007-9067-0

Bandura, A., Adams, N. E., Hardy, A. B., & Howells, G. N. (1980). Tests of the generality of self-efficacy theory. *Cognitive Therapy and Research, 4*, 39–66. https://doi.org/10.1007/BF01173354

Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior, 52*, 293–306. https://doi.org/10.1016/j.chb.2015.05.038

Baumgartner, S. E., Lemmens, J. S., Weeda, W. D., & Huizinga, M. (2017). Measuring media multitasking: Development of a short measure of media multitasking for adolescents. *Journal of Media Psychology: Theories, Methods, and Applications, 29*, 92–101. https://doi.org/10.1027/1864-1105/a000167

Beckmann, J. F. (2010). Taming a beast of burden – on some issues with the conceptualisation and operationalisation of cognitive load. *Learning and Instruction, 20*, 250–264. https://doi.org/10.1016/j.learninstruc.2009.02.024

Beckmann, J. F., & Goode, N. (2017). Missing the wood for the wrong trees: On the difficulty of defining the complexity of complex problem solving scenarios. *Journal of Intelligence, 5*, 15. https://doi.org/10.3390/jintelligence5020015

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research, 31*, 445–457. https://doi.org/10.1016/S0883-0355(99)00014-2

Boekaerts, M. (2017). Cognitive load and self-regulation: Attempts to build a bridge. *Learning and Instruction, 51*, 90–97. https://doi.org/10.1016/j.learninstruc.2017.07.001

Booker, C. L., Kelly, Y. J., & Sacker, A. (2018). Gender differences in the associations between age trends of social media interaction and well-being among 10-15 year olds in the UK. *BMC Public Health, 18*, 321. https://doi.org/10.1186/s12889-018-5220-4

Breems, N., & Basden, A. (2014). Understanding of computers and procrastination: A philosophical approach. *Computers in Human Behavior, 31*, 211–223. https://doi.org/10.1016/j.chb.2013.10.024

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*, 245–281. https://doi.org/10.3102/00346543065003245

Cerezo, R., Esteban, M., Sánchez-Santillán, M., & Núñez, J. C. (2017). Procrastinating behavior in computer-based learning environments to predict performance: A case study in moodle. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.01403

Cheng, S. L., & Xie, K. (2021). Why college students procrastinate in online courses: A self-regulated learning perspective. *The Internet and Higher Education, 50*, Article 100807. https://doi.org/10.1016/j.iheduc.2021.100807

Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers & Education: Artificial Intelligence, 1*, Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Chen, X., Zou, D., & Xie, H. (2022). A decade of learning analytics: Structural topic modeling based bibliometric analysis. *Education and Information Technologies*. https://doi.org/10.1007/s10639-022-11046-z

Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience, 17*, 718–731. https://doi.org/10.1038/nrn.2016.113

Cleary, T. J., & Callan, G. L. (2018). Assessing self-regulated learning using microanalytic methods. In J. A. Greene, & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (2 ed., pp. 338–351). New York: Routledge. https://doi.org/10.4324/9781315697048-22.

Cleary, T. J., Callan, G. L., & Zimmerman, B. J. (2012). Assessing self-regulation as a cyclical, context-specific phenomenon: Overview and analysis of SRL microanalytic protocols. *Education Research International 2012*, 1–19. https://doi.org/10.1155/2012/428639

Dessì, D., Cirrone, J., Recupero, D. R., & Shasha, D. (2018). Supernoder: A tool to discover over-represented modular structures in networks. *BMC Bioinformatics, 19*, 318. https://doi.org/10.1186/s12859-018-2350-8

Evans, T. M., Bira, L., Gastelum, J. B., Weiss, L. T., & Vanderford, N. L. (2018). Evidence for a mental health crisis in graduate education. *Nature Biotechnology, 36*, 282–284. https://doi.org/10.1038/nbt.4089

Eynon, R., Hjorth, I., Yasseri, T., & Gillani, N. (2016). Understanding communication patterns in MOOCs. In S. Elatia, D. Ipperciel, & O. Zaïane (Eds.), *Data mining and learning analytics* (pp. 207–221). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118998205.ch12.

Filvà, D. A., Forment, M. A., García-Peñalvo, F. J., Escudero, D. F., & Casañ, M. J. (2019). Clickstream for learning analytics to assess students' behavior with scratch. *Future Generation Computer Systems, 93*, 673–686. https://doi.org/10.1016/j.future.2018.10.057

Fournier-Viger, P., Gomariz, A., Campos, M., & Thomas, R. (2014). Fast vertical mining of sequential patterns using co-occurrence information. In V. S. Tseng, T. B. Ho, Z. H. Zhou, A. L. P. Chen, & H. Y. Kao (Eds.), *Advances in knowledge discovery and data mining* (pp. 40–52). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-06608-0_4.

Fournier-Viger, P., Lin, J. C. W., Gomariz, A., Gueniche, T., Soltani, A., Deng, Z., & Lam, H. T. (2016). The spmf open-source data mining library version 2. In B. Berendt, B. Bringmann, É. Fromont, G. Garriga, P. Miettinen, N. Tatti, & V. Tresp (Eds.), *Machine learning and knowledge discovery in databases* (pp. 36–40). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-46131-1_8.

Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology, 34*, 18–29. https://doi.org/10.1016/j.cedpsych.2008.05.006

Harrer, M., Apolinário-Hagen, J., Fritsche, L., Drüge, M., Krings, L., Beck, K., Salewski, C., Zarski, A. C., Lehr, D., Baumeister, H., & Ebert, D. D. (2019). Internet- and app-based stress intervention for distance-learning students with depressive symptoms: Protocol of a randomized controlled trial. *Frontiers in Psychiatry, 10*, 361. https://doi.org/10.3389/fpsyt.2019.00361

Herodotou, C., Rienties, B., Boroowa, A., Zdrahal, Z., & Hlosta, M. (2019). A large-scale implementation of predictive learning analytics in higher education: The teachers'role and perspective. *Educational Technology Research & Development, 67*, 1273–1306. https://doi.org/10.1007/s11423-019-09685-0

Hong, J. C., Lee, Y. F., & Ye, J. H. (2021). Procrastination predicts online self-regulated learning and online learning ineffectiveness during the coronavirus lockdown. *Personality and Individual Differences, 174*, Article 110673. https://doi.org/10.1016/j.paid.2021.110673

Hu, X., Luo, L., & Fleming, S. M. (2019). A role for metamemory in cognitive offloading. *Cognition, 193*, Article 104012. https://doi.org/10.1016/j.cognition.2019.104012

Inie, N., & Lungu, M. F. (2021). Aiki - turning online procrastination into microlearning. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–13). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3411764.3445202.

Junco, R. (2012). The relationship between frequency of facebook use, participation in facebook activities, and student engagement. *Computers & Education, 58*, 162–171. https://doi.org/10.1016/j.compedu.2011.08.004

Kaplan, A. (2008). Clarifying metacognition, self-regulation, and self-regulated learning: What's the purpose? *Educational Psychology Review, 20*, 477–484. https://doi.org/10.1007/s10648-008-9087-2

Kelley, W. M., Wagner, D. D., & Heatherton, T. F. (2015). In search of a human self-regulation system. *Annual Review of Neuroscience, 38*, 389–411. https://doi.org/10.1146/annurev-neuro-071013-014243

Küchler, A. M., Albus, P., Ebert, D. D., & Baumeister, H. (2019). Effectiveness of an internet-based intervention for procrastination in college students (studicare procrastination): Study protocol of a randomized controlled trial. *Internet Interventions, 17*, Article 100245. https://doi.org/10.1016/j.invent.2019.100245

Li, L., Gao, H., & Xu, Y. (2020). The mediating and buffering effect of academic self-efficacy on the relationship between smartphone addiction and academic procrastination. *Computers & Education, 159*, Article 104001. https://doi.org/10.1016/j.compedu.2020.104001

Liu, M., Wu, L., & Yao, S. (2016). Dose–response association of screen time-based sedentary behaviour in children and adolescents and depression: A meta-analysis of observational studies. *British Journal of Sports Medicine, 50*, 1252–1258. https://doi.org/10.1136/bjsports-2015-095084

Loeffler, S. N., Stumpp, J., Grund, S., Limberger, M. F., & Ebner-Priemer, U. W. (2019). Fostering self-regulation to overcome academic procrastination using interactive ambulatory assessment. *Learning and Individual Differences, 75*, Article 101760. https://doi.org/10.1016/j.lindif.2019.101760

Ludvigsen, S., Cress, U., Law, N., Stahl, G., & Rosé, C. P. (2018). Multiple forms of regulation and coordination across levels in educational settings. *International Journal of Computer-Supported Collaborative Learning, 13*, 1–6. https://doi.org/10.1007/s11412-018-9274-1

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from big data: Identifying self-regulated learning strategies in massive open online courses. *Computers in Human Behavior, 80*, 179–196. https://doi.org/10.1016/j.chb.2017.11.011

Meier, A., Reinecke, L., & Meltzer, C. E. (2016). "Facebocrastination"? predictors of using facebook for procrastination and its effects on students' well-being. *Computers in Human Behavior, 64*, 65–76. https://doi.org/10.1016/j.chb.2016.06.011

Moissa, B., Bonnin, G., & Boyer, A. (2019). Exploiting wearable technologies to measure and predict students' effort. In I. Buchem, R. Klamma, & F. Wild (Eds.), *Perspectives on wearable enhanced learning (WELL): Current trends, research, and practice* (pp. 411–431). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-64301-4_19.

Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., & Seppänen, T. (2019). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in Human Behavior, 100*, 298–304. https://doi.org/10.1016/j.chb.2018.12.019

Nussbaumer, A., Dahn, I., Kroop, S., Mikroyannidis, A., & Albert, D. (2015). Supporting self-regulated learning. In S. Kroop, A. Mikroyannidis, & M. Wolpers (Eds.), *Responsive open learning environments* (pp. 17–48). Cham: Springer. https://doi.org/10.1007/978-3-319-02399-1_2.

Osatuyi, B., & Turel, O. (2018). Tug of war between social self-regulation and habit: Explaining the experience of momentary social media addiction symptoms. *Computers in Human Behavior, 85*, 95–105. https://doi.org/10.1016/j.chb.2018.03.037

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00422

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom. *Journal of Educational Psychology, 82*, 33–40.

Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self-regulated learning. In G. Schraw, & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 43–97). Lincoln, NE: Buros Institute of Mental Measurements.

Pogorskiy, E., Beckmann, J. F., Joksimovic, S., Kovanovic, V., & West, R. (2018). Utilising a virtual learning assistant as a measurement and intervention tool for self-regulation in learning. In *2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE), IEEE* (pp. 846–849). https://doi.org/10.1109/TALE.2018.8615130

Rasheed, R. A., Kamsin, A., & Abdullah, N. A. (2020). Challenges in the online component of blended learning: A systematic review. *Computers & Education, 144*, Article 103701. https://doi.org/10.1016/j.compedu.2019.103701

Reich, J., & Ruipérez-Valiente, J. A. (2019). The MOOC pivot. *Science, 363*, 130–131. https://doi.org/10.1126/science.aav7958

Reparaz, C., Aznárez-Sanado, M., & Mendoza, G. (2020). Self-regulation of learning and MOOC retention. *Computers in Human Behavior, 111*, Article 106423. https://doi.org/10.1016/j.chb.2020.106423

Rizvi, S., Rienties, B., Rogaten, J., & Kizilcec, R. F. (2020). Investigating variation in learning processes in a FutureLearn MOOC. *Journal of Computing in Higher Education, 32*, 162–181. https://doi.org/10.1007/s12528-019-09231-0

Robal, T., Zhao, Y., Lofi, C., & Hauff, C. (2018). IntelliEye: Enhancing MOOC learners' video watching experience through real-time attention tracking. In *Proceedings of the 29th on hypertext and social media* (pp. 106–114). New York, NY, USA: ACM. https://doi.org/10.1145/3209542.3209547.

le Roux, D. B., Parry, D. A., Totolo, A., Iyawa, G., Holloway, J., Prenter, A., & Botha, L. (2021). Media multitasking, online vigilance and academic performance among students in three southern african countries. *Computers & Education, 160*, Article 104056. https://doi.org/10.1016/j.compedu.2020.104056

Siadaty, M., Gašević, D., & Hatala, M. (2016). Trace-based microanalytic measurement of self-regulated learning processes. *Journal of Learning Analytics, 3*, 183–214. https://doi.org/10.18608/jla.2016.31.11

Steel, P., Svartdal, F., Thundiyil, T., & Brothen, T. (2018). Examining procrastination across multiple goal stages: A longitudinal study of temporal motivation theory. *Frontiers in Psychology, 9*, 327. https://doi.org/10.3389/fpsyg.2018.00327

Sternberg, N., Luria, R., Chandhok, S., Vickers, B., Kross, E., & Sheppes, G. (2020). When facebook and finals collide - procrastinatory social media usage predicts enhanced anxiety. *Computers in Human Behavior, 109*, Article 106358. https://doi.org/10.1016/j.chb.2020.106358

Suchert, V., Hanewinkel, R., & Isensee, B. (2015). Sedentary behavior and indicators of mental health in school-aged children and adolescents: A systematic review. *Preventive Medicine, 76*, 48–57. https://doi.org/10.1016/j.ypmed.2015.03.026

Svartdal, F., Dahl, T. I., Gamst-Klaussen, T., Koppenborg, M., & Klingsieck, K. B. (2020). How study environments foster academic procrastination: Overview and recommendations. *Frontiers in Psychology, 11*, 3005. https://doi.org/10.3389/fpsyg.2020.540910

Tremblay, M. S., LeBlanc, A. G., Kho, M. E., Saunders, T. J., Larouche, R., Colley, R. C., Goldfield, G., & Gorber, S. (2011). Systematic review of sedentary behaviour and health indicators in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity, 8*, 98. https://doi.org/10.1186/1479-5868-8-98

Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education, 122*, 119–135. https://doi.org/10.1016/j.compedu.2018.03.018

Wessel, J., Bradley, G. L., & Hood, M. (2019). Comparing effects of active and passive procrastination: A field study of behavioral delay. *Personality and Individual Differences, 139*, 152–157. https://doi.org/10.1016/j.paid.2018.11.020

Winne, P. H. (2017). Cognition and metacognition within self-regulated learning. In P. A. Alexander, D. H. Schunk, & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2 ed., pp. 36–48). Routledge. https://doi.org/10.4324/9781315697048-3.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Erlbaum.

Yun, J. T., Vance, N., Wang, C., Marini, L., Troy, J., Donelson, C., Chin, C. L., & Henderson, M. D. (2020). The social media macroscope: A science gateway for research using social media data. *Future Generation Computer Systems, 111*, 819–828. https://doi.org/10.1016/j.future.2019.10.029

Ziegler, N., & Opdenakker, M. C. (2018). The development of academic procrastination in first-year secondary education students: The link with metacognitive self-regulation, self-efficacy, and effort regulation. *Learning and Individual Differences, 64*, 71–82. https://doi.org/10.1016/j.lindif.2018.04.009

Zimmerman, B. J. (2000). Attaining self-regulation. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*, 166–183. https://doi.org/10.3102/0002831207312909