# Designing Visual and Interactive Self-Monitoring Interventions to Facilitate Learning: Insights from Informal Learners and Experts

Rimika Chaudhury, and Parmit K. Chilana

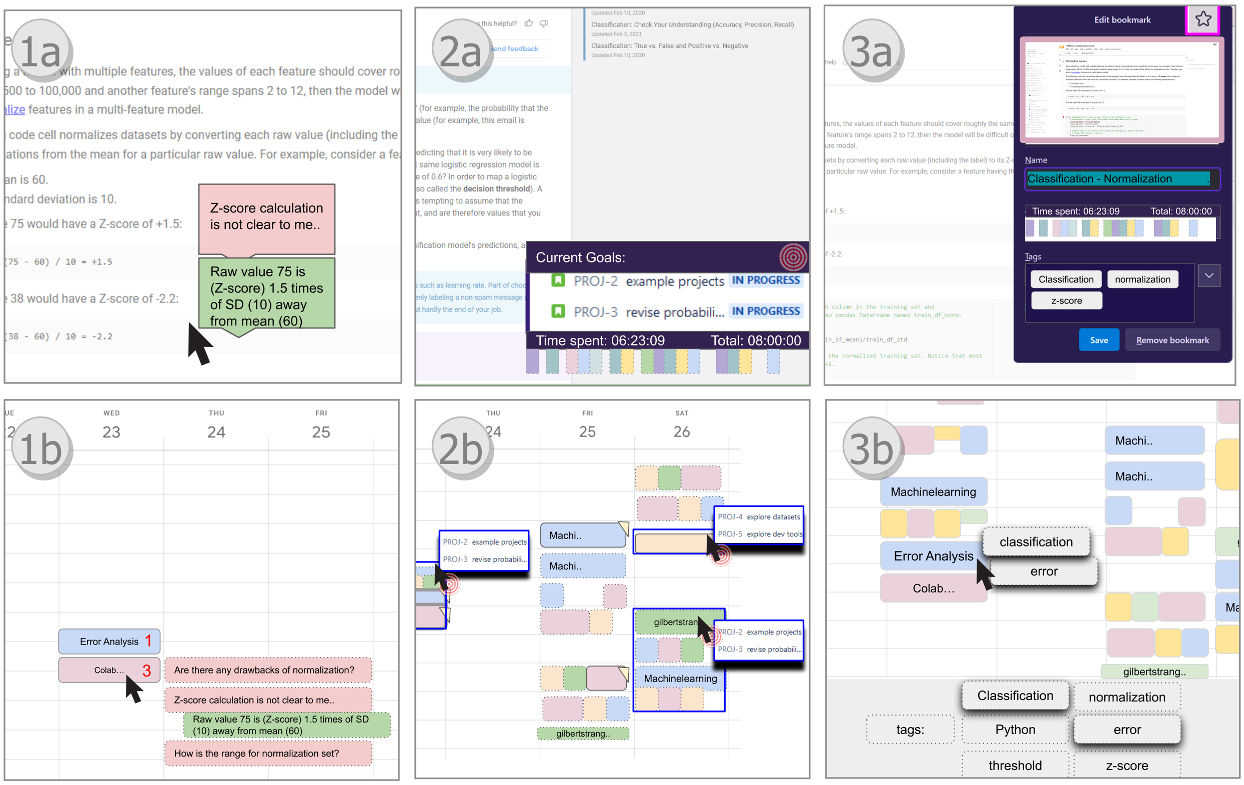


Fig. 1: Three examples of paper-based visual overviews used in Study 1. Design idea #1: The muddy point approach is illustrated by Figures 1a. and 1b. Design idea #2 Goal-setting is shown in 2a. and 2b. The third column shows design idea #3 Vocabulary-based filter.

**Abstract**—Informal learners of computational skills often find it difficult to self-direct their learning pursuits, which may be spread across different mediums and study sessions. Inspired by self-monitoring interventions from domains such as health and productivity, we investigate key requirements for helping informal learners better self-reflect on their learning experiences. We carried out two elicitation studies with paper-based and interactive probes to explore a range of manual, automatic, and semi-automatic design approaches for capturing and presenting a learner's data. We found that although automatically generated visual overviews of learning histories are initially promising for increasing awareness, learners prefer having controls to manipulate overviews through personally relevant filtering options to better reflect on their past, plan for future sessions, and communicate with others for feedback. To validate our findings and expand our understanding of designing self-monitoring tools for use in real settings, we gathered further insights from experts, who shed light on factors to consider in terms of data collection techniques, designing for reflections, and carrying out field studies. Our findings have several implications for designing learner-centered self-monitoring interventions that can be both useful and engaging for informal learners.

**Index Terms**—Learner-centered Design, Informal Learners, Elicitation Study, Expert Validation, Self-monitoring Techniques

## 1 Introduction

Millions of people around the world are turning to informal learning resources online to develop computational skills [2] and to keep up with the demands of remote work and learning [45]. Informal learners can access a variety of educational content in different formats (e.g., articles, videos, forums, e-books) and pursue their learning at their own pace. However, these learners can face a number of barriers in their informal learning pursuits [9, 13, 68], particularly in self-monitoring their progress [11].

One of the key challenges that informal learners face in self-directing [39] their learning is that they often lack awareness of their own strate-

- *Rimika Chaudhury is with Simon Fraser University. E-mail: rimika_chaudhury@sfu.ca.*
- *Parmit K. Chilana is with Simon Fraser University. E-mail: pchilana@cs.sfu.ca.*

*Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx*

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2024.3366469

gies and potentially unhelpful behaviors. For example, they may rely on suboptimal trial and error [11, 18] and oscillate between different media and resources without a systematic strategy. Furthermore, the feedback that a learner receives organically in a social setting, such as a classroom, is missing in informal learning settings where the onus is on the learner to monitor their own progress and assess comprehension.

Since activities of self-reflection have long been shown to be useful in formal classroom-based educational contexts [39], we wondered how these activities could be designed for informal learners who largely rely on online resources and pursue their learning individually. For example, *what if learners could observe their learning patterns across different media and study sessions? What if learners could monitor their time spent and reflect on their trial-and-error behaviors to better self-direct their efforts?*

In our research, we take a learner-centered design-oriented [62] approach to explore ways of allowing learners to tap into their own learning experiences for self-reflection. To explore this design space, we take inspiration from prior work in self-monitoring and self-tracking in domains such as health, well-being, and productivity [15, 48], which has demonstrated several benefits of tracking progress using manual to automatic approaches. In the context of informal learning, recording and reflecting on learning activities could also help raise self-awareness at various stages of learning and the learner's overall success [42]. However, recent research suggests that existing tracking tools offer little flexibility with data collection and presentation nor support scaffolding for goal-setting, which may limit the opportunities for user-driven self-reflection [14]. Thus, we explore the design space of self-monitoring interventions for informal learners, emphasizing goal-setting and reflection on learning progress.

In this paper, we use a two-part elicitation study to synthesize requirements for the design of self-monitoring tools and techniques for informal learners of computational skills. Using the design probes approach [67], we adapted features from existing tracking tools [15, 37, 48] to explore a range of ideas between two extremes: completely manual methods where the learner is deeply involved in both data collection and presentation [1, 20], and automated techniques where the learner's involvement is minimal [36]. Most of the ideas that we explored leverage *semi-automated* ways of self-monitoring [15].

In our first study, we showed 8 participants our design probes in the form of paper-based mock-ups to elicit their perceptions of what kind of data is useful to reflect on and understand the extent to which learners may want to be involved in self-monitoring activities. We learned that visual overviews of learning histories can be insightful for learners and provide them with a way to meaningfully engage with their data. However, we were still unclear about how learners may use any of this information for reflecting on their learning and goal-setting as our paper-based mock-ups did not capture the dynamics of interaction for reflective activities and planning.

In our second study, we used the insights from our first study to design three interactive probes that varied the presentation of the overviews along three dimensions: temporal, resource-type, and topic. Our second elicitation study with 12 participants focused on understanding how learners would use the different interactive overviews to evaluate past efforts and plan the next steps. We found that semi-automatic approaches appealed to our participants the most, consistent with other studies [15]. Participants appreciated the at-a-glance summary provided by the broad categorizations of resources and welcomed the granular breakdown of daily activities with an integrated to-do list (e.g., in the Temporal Overview) to reflect on multiple goal-pursuits. Most participants felt hesitant about the completely automatic generation of subtopic clusters in the Topic-based Overview for the lack of transparency and control it offered. Participants expressed eagerness to be involved in actions (e.g., tagging, annotating) that could improve the utility of the overviews for planning and sharing.

Although our probe-based studies offered valuable insights into learners' self-monitoring needs, we considered diversifying our perspectives to interpret our findings by drawing on the perspectives of six HCI and visualization experts in the field of self-monitoring. In our third study, we aimed to gain insights into factors influencing the design of tools for collecting personal data in real-world scenarios, as these factors could significantly impact users. We consulted experts whose work inspired ours and included specialists in self-monitoring research across various domains like health, well-being, and productivity. Experts emphasized the importance of informed participant onboarding regarding data collection and privacy measures. They also highlighted the design tension between user flexibility and structured guidance for personal data-based reflection. Additionally, experts cautioned against potential negative self-perceptions that individuals may harbor about self-monitoring and advised designing interventions to highlight events with positive associations in the self-monitored data.

The key results from our first two studies contribute an initial set of requirements for designing self-monitoring interventions for informal learners. Adopting a user-centered approach, our work carefully examines how learners interact with various visual data presentations, prioritizing their preferences and values over automatic visualizations. The last study further integrates insights from experts in HCI and visualization, contributing practical insights into building and deploying self-monitoring tools. Our qualitative and design-based approach, involving multiple user groups and methods, triangulates and complements other survey-based and experimental-based methods [12, 54, 57, 58, 66], offering deeper insights into user engagement with self-monitoring tools. We make several design recommendations for promoting self-reflection to help learners leverage their learning histories and be more engaged and in control of their informal learning pursuits.

## 2 Related Work

Our work builds upon research on supporting informal and self-directed learning of computation skills, visualization techniques used in self-monitoring interventions in HCI, and data-driven visualizations for supporting reflections in learning.To organize the relevant literature, we used the dimensions outlined for personal visualization and personal visual analytics systems [28], focusing on 1) data collection (data scope, effort, agency), 2) insight generation (interactive and automated data presentation), and, 3) user involvement in design and evaluation.

### 2.1 Supporting Informal and Self-directed Learning

Learning support tools for self-monitoring and self-regulation, such as goal-setting and progress reports, are often found on dedicated platforms for large-scale course delivery such as Massive Open Online Courses (MOOCs), and CS1/CS2 university courses [21, 22, 26]. More recently, explorations in alternative mediums, such as immersive technology, are investigating the impact of Augmented Reality (AR) and Virtual Reality (VR) on collaborative learning and knowledge gain, particularly for content that is challenging to observe such as physical or environmental phenomena [61]. However, these works focus on the creation and implementation of immersive learning environments, emphasizing practical use guidelines and cognitive involvement [57, 58]. While these works align with our emphasis on designing innovative visual learning approaches, our work distinguishes itself by shifting the focus from structured teacher-led classrooms to less formal learner-led situations.

We describe *informal learning* as a self-directed learning pursuit where the learner takes the initiative to set their learning objectives and utilize diverse online resources, occasionally seeking support from mentors or peers for feedback. Recent HCI studies highlight a growing population of informal learners interested in advanced computational skills [11, 13, 25, 68]. These studies show that informal learners face challenges distinct from those observed in formal settings arising from factors such as programming environments [40], quality of lessons [25], and the learners' own assumptions and biases [11, 55]. A consistently observed theme is that informal learners often rely on exploration and trial-and-error strategies in resource selection and implementation, lacking opportunities for reflective thinking [11, 18, 44]. While interventions for self-reflection in formal programming courses are emerging [42, 56], their applicability in informal, less-structured environments remains unclear [11]. Our contribution lies in exploring the design of self-monitoring interventions and related visualizations for

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2024.3366469

informal learners and understanding the types of data and interactions that could benefit their self-directed learning.

### 2.2 Visualization-based Self-Monitoring Tools in HCI

Visualizations have emerged as a key feedback method in user-driven approaches for self-monitoring. Static visualizations offer quick progress overviews, while interactive tools allow users to explore more complex insights using longitudinal, time-series graphs with contextual factors or multiple visualizations simultaneously. Fitness, sleep, and food tracking apps integrate diverse data, enabling users to identify patterns and trends in their history, conduct self-experimentation, and explore lifestyle influences on various aspects [14]. Explorations in visual analytics and immersive technologies are investigating ways to enhance interaction and insight generation by drawing on quantified data and automated analyses [29, 41, 72]. These tools aim to facilitate dialogic reflection, encouraging the exploration of data by presenting correlations and potential causality. However, these works lack exploration of human-computer interaction aspects of data collection - limiting support for recording and reflecting on open-ended user accounts [17]. In contrast, our work investigates how automated systems can support learners in data collection, allowing them to define scope, control involvement, and contribute to and curate collected data [28].

Previous research on self-monitoring interventions, particularly in health, well-being, and productivity, highlights the positive impact of self-tracking practices on self-awareness and self-reflection [33, 35]. However, the burden of manual data collection through texts or photos has been identified as a potential obstacle for long-term engagement [15, 20]. Attempts to address this challenge through automatic data capture have shown that a lack of user engagement significantly diminishes awareness and involvement in the tracking practice [16, 38, 47, 50, 64]. Conversely, non-digitized, manual tracking styles reveal that deliberate involvement with personal data can promote self-reflection and self-expression [1, 43]. In response, researchers are exploring semi-automated approaches that combine automatic and manual methods, emphasizing the importance of understanding users' motivation and context [15, 37, 70]. Semi-automated approaches can vary in the extent of automation, ranging from extensive automation that relegates minimal tasks to users (e.g., correcting collected data) or provides minimal machine assistance (e.g., generating time stamps)

Drawing inspiration from these interventions, we sought to understand what data informal learners care about and how they prefer to explore and interact with such information. We synthesize these insights to derive design requirements for a semi-automated self-monitoring tool intended to support self-directed learning.

### 2.3 Data-Driven Approaches and Visualizations for Supporting Learning Reflections

In Education and Learning Sciences, researchers are increasingly turning to data-driven and visualization methods to enhance online learning experiences [54]. Visualizations play a significant role in self-monitoring learning, notably through Learning Analytics (LA), with Learning Analytics Dashboards (LAD) being a common tool for presenting learning traces derived from LA data. These traces include logs of user activity, artifacts created by learners, and test results presented through various static visualization types like bar charts, line graphs, and tables [12]. However, current feedback mechanisms in LADs primarily cater to instructors and administrators, leaving out considerations for self-directed learners [6, 23, 32]. There is a growing emphasis on empowering learners to explore and analyze their data, encouraging goal-setting, self-evaluation, and reflections on assessments [3, 26, 59, 69]. Yet, existing learning analytics tools may not align with individual learning needs, hindering opportunities for self-exploration and discovery within learners' contexts. The challenge then lies in creating tools that not only leverage data but also cater to the diverse learning approaches of self-directed learners [3, 66].

Moreover, a notable gap exists in guiding the design of self-monitoring data-driven tools for learning. Existing approaches have relied on methods like usability studies or expert-led cognitive walkthroughs on refined systems [12, 54]. Prior research in this field has minimally involved users, typically only in the evaluation stage and seldom during the early design phase. This approach provides a limited understanding of how learners perceive and utilize these tools and what data they find useful in real contexts. To address this, our work adopts a learner-centered, iterative design approach, aiming to understand the types of data and interactions beneficial for fostering reflection and supporting learners. This approach aligns with the growing trend of employing learner-centered and design-based methods [62, 67] in the broader visualization community, allowing us to uncover learners' contexts and needs, informing the design of visualizations tailored for self-monitoring in online informal learning experiences [3, 27, 32].

## 3 Study 1: Eliciting Requirements Using Paper-Based Mockups

Guided by learner-centered design [62], our research addresses the diversity in learners' objectives, motivations, and challenges. We aimed to understand informal learners' perceptions of self-monitoring techniques and identify key considerations for designing self-directed learning tools. Taking a qualitative design-probe approach [67], we adapted self-monitoring intervention attributes from domains like health and productivity (see Table 1). In Study 1, we evaluated learners' perceptions of manual, automatic, and semi-automatic self-monitoring techniques in informal learning, exploring what types of data are perceived to be useful for self-reflection and the extent of learners' willingness to engage in data collection and presentation tasks. We explored design interventions using paper-based mockups (simplified sketches of visual overviews) as they allowed us to illustrate a wider range of ideas and prompted participants to speculate on the utility of seeing their personal data and learning patterns.

### 3.1 Exploring Visual Overviews using Paper Mockups

To create the six paper-based mock-ups (see Fig. 1), we considered methods for data recording, presentation, and motivational components, such as goal-setting and prototyping materials (see Table 1). Our designs encompassed a spectrum from automated tracking with limited user control (design idea #3) to manual recording methods (design idea #6). Presentation layouts included a weekly summary with daily and hourly details, adapted from existing models [49, 50], and tailored to specific learning scenarios. Additionally, we explored pre-session goal-setting and in-session reflective prompts for goal alignment (design idea #2). We probed about material preferences by including mockups suggesting physical (design idea #6) and digital tracking (design ideas #1 - #5).

**Design idea #1 - Muddy Point:** Grounded in learning sciences research [31], this design investigates the effectiveness of externalizing and visualizing both resolved and unresolved questions to enhance reflections on challenging concepts or "muddy points." The probe illustrates the interaction of annotating questions and answers within resources and presents this information in a calendar-based overview alongside the relevant resource (see Fig. 1.1a and 1.1b).

**Design idea #2 - Goal Setting:** This design utilizes a Kanban-inspired board, commonly employed in software management [49], to define and track goals. The probe introduces the concept of maintaining goal orientation through pop-up interactions for a manual association of pre-set goals with newly visited web pages during study sessions (see Fig. 1.2a). The overview demonstrates how resources can be grouped based on goal associations, facilitating reflections on goal alignment at the end of the learning session (see Fig. 1.2b).

**Design idea #3 - Vocabulary-based Filter:** Since learners initially have a limited vocabulary in a new learning domain [24], we explored the idea of automatic generation of keywords based on content viewed and saved by learners (see Fig. 1.3a). The design presents a weekly collection of keywords that can be used as filters to scan for related saved resources in the overview itself (see Fig. 1.3b). This idea aimed to gather insights on completely automated support for collecting data regarding newly learned vocabulary, and how they could be used for reflecting on resources.

**Design idea #4 - Predicting Usefulness:** Informal learners often face challenges in decision-making, potentially due to lower competency

Table 1: Four design attributes that were considered in the design of the paper-based mock-ups, based on the literature of self-tracking tool designs in productivity, and health and well-being domains.

| |
|---|
| **1. Data recording:** Automatic tracking through sensors or logs may reduce the burden of capture [15], whereas manual tracking allows users to be more aware and engaged with data [4]. Leveraging both forms of data-collection through semi-automated approaches provides the flexibility of shifting the control between the user and the system as appropriate [15] |
| **2. Data presentation:** The layout of the data should allow users to glean insights through exploration and therefore should provide glanceable summaries, as well as allow the users to manipulate the visualization to discover details on demand. Such goals can be supported through interactions like selection, filtering, and zooming [19,71] |
| **3. Motivational components:** Self-monitoring tools are often designed with behavior change as a target outcome. Such goals are often supported by persuasive techniques such as reminders, nudges, and prompts, for goal-setting and goal-adherence [30]. |
| **4. Material considerations:** Physical materiality of tracking tools often allow for mindful, slow-paced explorations and self-reflection [1,65]. On the other hand, digital tools can afford long-term tracking, ease data collection overhead, and offer powerful interactions for exploring data [52]. |

with new content [51]. Reflecting on the utility of a resource may aid learners in making more informed choices [10]. This design introduces interactions to explore participants' perceptions of making explicit judgments about visited web pages to reflect on their usefulness. The overview includes a browser-based form, enabling users to insert notes, express emotions about the resource, and indicate useful subsections within resources through check-mark annotations.

**Design idea #5 - Cross-Referencing:** Informal learners often assimilate information from diverse sources, as establishing connections between these sources enhances comprehension and reflection [60]. This design investigates the usefulness of visualizing cross-referencing as learning-activity data. The mock-up depicts a webpage where a link to an externally bookmarked page from a prior session can be added. The calendar-based overview indicates which resources contain cross-references and incorporates a chatbot for automatic analysis of visited web pages to gather insights on learners' sentiments about automatic summaries of their browsing activities.

**Design idea #6 - Bullet Journals**: This mock-up explores perceptions of manual and physical ways of tracking learning using two pages of a pocket-sized diary. Despite offering some openness, the small size and bullet-point approach [5] create boundedness. The mock-up data includes dates, resource titles, and comments or questions logged by the learner in the scenario. The design enhances flexibility by employing symbols, like question marks for points of confusion and arrows to indicate scheduling tasks for future dates.

### 3.2 Study Procedure and Analysis

Participants completed a brief questionnaire covering demographics, their definition of progress, and success in a learning context. We used each paper-based mockup as a conversation starter, contextualized within an informal learning scenario. We next asked the participants to evaluate the pros and cons of each idea in facilitating reflections on their learning, considering resource quality, time spent, and progress. We conducted in-person interviews lasting approximately an hour, and participants received CAD 20 gift cards for their participation.

**Participants:** Our goal was to recruit participants who were learning complex technical skills using informal online resources. We reached out to the contacts of the research team and to others using snowball sampling and through the university mailing list. Eight participants (4M/4F) who were all students (6 graduate and 2 undergraduate students) between 19-35 years of age signed up for our study. Our participants reported self-learning technical skills such as machine learning (ML), web development, and data visualization.

**Analysis:** The interviews were audio recorded with the participants' consent and later transcribed. Two researchers were involved in analyzing the qualitative data from the interviews. We used an inductive analysis approach [63], beginning with open coding to inspect each transcript. While coding, we considered how the responses were related to participants' perceptions of meaningful data for self-reflection and their perceptions on data collection methods. We assigned multiple codes where necessary and had regular discussions with the research team to reconcile our final coding scheme. We performed axial coding to explore themes around our research questions and synthesized key insights as our results.

### 3.3 Key Findings from Paper-Based Elicitation

All participants found the idea of self-monitoring to be useful for gauging learning progress and leaned towards semi-automatic approaches to tracking. All but one (7/8) participant explained that while automatic methods could provide a convenient and systematic way to record their learning attempts, having manual control over certain aspects of data collection and presentation could improve the utility of visual overviews. Most participants (6/8) wanted to be able to indicate to the system what kind of information to record, such as the frequency of concepts marked as relevant, difficult, urgent, or important. Additionally, participants wanted to be able to add, remove, or edit data in the overviews to make them more accurate, specific, or useful for planning and prioritizing. Lastly, participants shared some concerns regarding goal-setting and gauging progress.

#### 3.3.1 Overviews create awareness of learning processes

Participants expressed that visual overviews could help them identify where they were spending time and assess the extent of their progress during the period. Most of the participants (6/8) found the hourly breakdown of daily activities a useful indicator of productivity and performance. The participants shared that the automatically generated visual overviews (see Fig. 1.1b, 1.2b, 1.3b) could additionally serve as a reminder of the topics within resources that they had viewed, and tasks they had accomplished during the week. Participants (7/8) acknowledged that visual overviews that make it possible to zoom in and explore specific data points, as opposed to aggregated data, could facilitate scanning and searching for useful resources and could help them save time while resuming a subsequent study session. Half of the participants (4/8) were in the habit of collecting useful links on note-taking apps for future access and mentioned that the overviews could serve as gateways to their personally curated collections of resources. P04 added that easier recognition of relevant resources through meaningful categorizations (such as the one used in the goal-setting probe) could also facilitate recollection of helpful strategies for a future task: *"...because you are clearly grouping everything, it's easier to look back at later.. I do open up old projects if I am trying to remember how I did something before."*

While all participants agreed that the time-based overviews could help them filter information based on days and hours, they also imagined other ways to parse their learning histories. Three participants said that they cared less about the exact time spent on a resource or activity and were more interested in gaining an overall idea about their engagement. Most participants (6/8) were interested in knowing how many novel concepts they had learned during a given study session (see Fig. 1.3b), or gleaning the topics they had been studying during the specified time (one week, in our study).

#### 3.3.2 Semi-automatic data-recording for purposeful revisits

The majority of participants (7/8) indicated that they were skeptical of the automatic identification of topics from the visited links and preferred to have some control over the topics presented in the overviews. While participants preferred minimal engagement with data collection processes during the study session, they were willing to fine-tune and organize the automatically detected topics to better indicate on the overviews which concepts they had studied and found relevant.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2024.3366469

Many participants (5/8) were apprehensive about seeing every visited link (and search times) on the overview, as much of this effort could have been wasteful. They wanted the overviews to mostly serve as a way to look back on fruitful pursuits and only wished to see resources with which they had meaningful interactions. These could be resources that they have accessed repeatedly, annotated with highlights, comments, and questions, or saved. In P01's view: *"[Progress] is not about the time we spend.. it is about having some sort of questions in mind. If [...] the questions are answered, then I made progress."*

Almost all participants (7/8) mentioned that overviews should present learning activities that could improve the value of a resource. Participants mentioned examples like classifying resources into broad categories based on topics or content type (e.g., segregating code tutorials and conceptual content), identifying relevant segments within articles or videos, and associating tasks with resources. Participants also saw value in using visual overviews to indicate the level of perceived difficulty and relevance of resources to current learning interests and prioritize resources. P04 noted that despite being annotated as useful, resources may still end up being difficult to revisit: *"you need to take action on them later for [annotations] to have value"*. In the next subsection, we touch on the potential of using overviews to help learners evaluate progress and identify actionable items.

### 3.3.3 In-context to-do lists for gauging learning progress

The majority of the participants (7/8) mentioned specific ways in which visual overviews could help them achieve a sense of progress. For example, five participants mentioned how a to-do list showing an account of completed and pending tasks and the provision to strike out tasks could help them assess their progress. However, four participants expressed concerns about using goal-oriented interventions (design idea #4), stating that initial goals may be *"super vague or sometimes incorrect"*(P07) and tend to evolve over time.

Participants expressed some hesitation in writing down goals before a learning session, as they usually figured it out "on the go". P03 shared how he tended to submit his code in *"one large commit on GitHub"* instead of thinking about it in "*chunks*" or work done to meet sub-goals. However, he explained how grouping resources based on the learners' areas of interest or projects could be helpful (see Fig. 1.2):

> *It's almost like you're giving this resource a metric because you've already given a metric in your head, you come into this resource with a bias, which is a good thing...you're able to reflect on yourself and on this resource, whether it's useful, or not.* (P03)

P06 also shared how the overviews could provide an objective measure of progress in terms of the number of novel topics covered (see Fig. 1.3), and could be useful for sharing for feedback as it is *"relatable to others"*. For more personal measures of progress, participants indicated that they were interested in seeing the "extent" of progress in the direction of a larger goal (e.g., a summary of consumed content, or an account of answered questions). P05 mentioned that partial progress meant completing *"75% of the web page"* or trying out *"example codes from tutorials and copy[ing] line by line [to] see what happens."*

From our paper-based study, we learned that participants wanted to improve performance at certain repetitive tasks (e.g. refining the output of a model) or determine if they had been successful in adapting existing examples to their specific needs. However, the paper-based mockups were limited in that they did not capture the dynamics of possible interactions learners may want to carry out while exploring overviews for reflections and planning. We wanted to probe more into what would make reflections more engaging and useful while using different types of visual interactive overviews.

## 4 Study 2: Exploring the Designs of Interactive Overviews

In Study 2, we refined the requirements from Study 1, by investigating how learners perceive interactive visual overviews as a means of reflecting on their learning patterns and gauging progress. We designed three interactive probes (see Fig. 2, Fig. 3, and Fig. 4) that highlighted different aspects, including time spent, resource types used, and topics covered in a learning session. As *at-a-glance summaries* are shown to be useful for providing quick insights on progress [19], we varied the data presentation to provide learners with different types of insights into their learning experiences.

Our research question was: *How do learners make use of interactive learning overviews that vary in presentation format (e.g., time-based, topic-based, resource-based) to reflect on recent learning efforts and to plan the next steps?* Next, we describe the design of the three interactive probes and the motivations behind each design.

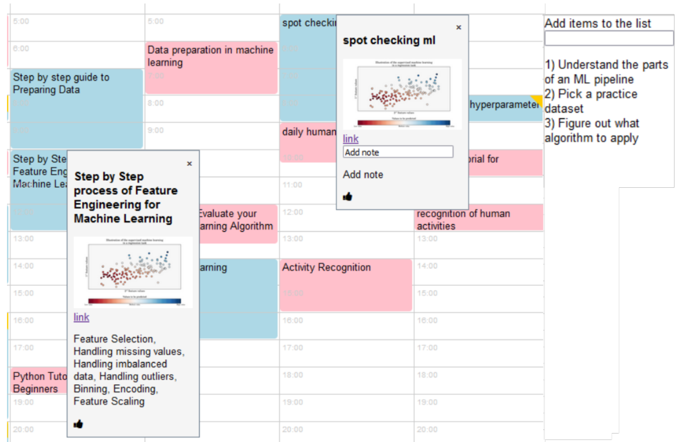


Fig. 2: *Temporal Overview* uses a calendar-based time-boxed layout to represent visited and annotated (red boxes) resources, with details of each resource displayed on demand. Also, includes a basic to-do list.

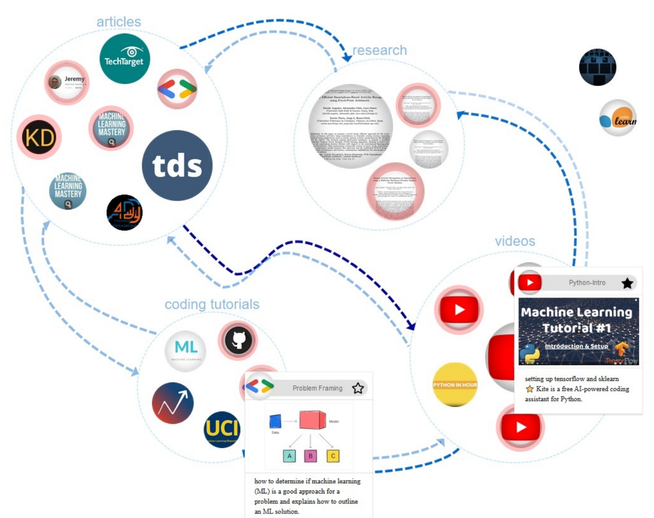


Fig. 3: *Resource-based Overview* represents resources grouped by resource-type (e.g., videos, tutorials). The arrows represent cumulative movements from one resource-type to another, with more frequent movements shown in a darker shade. To-do list is not shown in this figure.

### 4.1 Temporal Overview

**Motivation:** Our participants perceived the calendar-based overviews of learning resources to be useful for developing an awareness of productivity and have the potential for optimizing the time to look up saved resources. We wanted to further disentangle the benefits and drawbacks of temporal overviews.

**Description:** To help users track their learning and prompt reflection, the Temporal Overview (see Fig. 2) uses timeboxes [50] to show a weekly spread of the duration (hours) spent on each resource.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2024.3366469

Timeboxes are colored red when annotated or blue when unannotated. The timeboxes display a yellow dog-ear bookmark when a resource is judged and marked as important by the learner. We included keyword-based filtering based on the 9 most relevant keywords from the contents of the selected webpages to encourage reflection of newly learned concepts. Clicking on timeboxes displays resource details, such as snapshots, notes, and usefulness judgments (thumbs-up icon). This overview includes a basic in-context to-do list.

### 4.2 Resource-based Overview

**Motivation:** Our first study revealed that helping learners recognize different types of resources, such as segregating code tutorials from conceptual content, could prompt reflection. In addition, we wanted to investigate whether providing information about the sequence of resource access could also encourage reflective thinking.

**Description:** The Resource-based Overview (see Fig. 3) displays the learning medium (e.g., videos, articles, tutorials, forums, publications) and encourages reflection on content preferences. We wanted to probe the importance of resource titles for interpreting overviews and ways to simplify time-spent information. Resource circles, varying in size and displaying a logo of the source, represent each resource, with the title available on-demand through mouse-hover. Three circle sizes represent the relative time spent per resource. Arrows indicate the sequence of resource access, with darker shades of blue indicating a higher frequency of movement. Red borders indicate annotated or questioned resources. Details are available on-demand using a details-card per resource approach. This overview also includes an in-context to-do list, similar to the Temporal Overview.

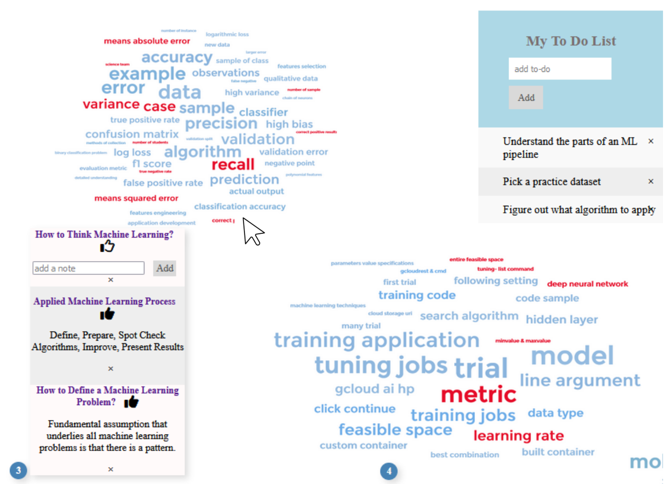


Fig. 4: *Topic-based Overview* clusters resources based on similarity of topics; word clusters show relevant keywords from each collection. Collection of resources are available in a list, on demand.

### 4.3 Topic-based Overview

**Motivation:** Based on Study 1 findings, we investigated the potential of topic or keyword-based word clusters in promoting self-reflection. Participants indicated that an account of newly learned content, such as keywords and concepts, could provide a sense of progress.

**Description:** The Topic-based Overview (see Fig. 4) presents word clusters that coalesce relevant keywords found within saved or annotated web pages based on the similarity of topics. For example, resources related to different topics, such as algorithms, introductory statistical concepts, advanced training and tuning-related concepts, and toolkits, are separated into different clusters. However, the criteria for such grouping are left to the participants to interpret from the words in each cluster. Red highlighted words suggest outstanding questions related to the corresponding topic. Word size implies the relative time spent on the underlying resource and topic. The source webpages for each cluster are available on demand as a list, with personal notes and judgments of usefulness. Consistent with the other probes, the design includes an in-context to-do list.

### 4.4 Implementation of Interactive Probes

All of our interactive probes were semi-automatic by design as they showed the potential to capture certain data automatically (e.g., the time of visit, titles, snapshots, and resource URLs), but included opportunities for users to manually provide additional details (e.g., assessment of resource usefulness, annotations, notes, and to-do items). Each overview also included some capacity for edits (e.g., removing or adding notes and resources), which we used to explore perspectives on editing for personal use versus sharing with one other person for feedback. We used the same set of mocked-up data and shared characteristics to design each visual overview. The overviews summarized the past week, showed learner interactions with resources (e.g., annotated questions, bookmarks), and included a to-do list. Basic interactions were available, such as filtering, selecting items, and navigating to external links. The in-context to-do list was designed to be simple, encouraging participants to share their perspectives on additional useful actions for reflection and planning. The probes were implemented as high-fidelity prototypes using JavaScript frameworks and libraries (React and JQuery), along with wireframing toolkits (Axure RP).

### 4.5 Study Procedure and Analysis

For our second elicitation study, we conducted semi-structured interviews using interactive probes to facilitate discussion. We first introduced participants to a scenario where a learner, Jay, uses online informal resources online to learn ML. Jay's learning activities online (e.g., visited websites, annotations, time) were illustrated in our probes as visual overviews. The data used to create the overviews were mocked up by the researchers. We asked our participants to assume that such overviews could be generated automatically in real-time as they proceeded through their learning and reminded them that the overviews were only partial representations of the learning activities for one week. While the general structure of the interview questions remained the same as in Study 1, we refined certain questions to probe more deeply about specific design attributes using the interactive probes. For example, we asked our participants to interact with each probe and share with us how they may use such visual overviews to reflect on their past learning sessions, plan ahead for future sessions, and how they perceived the pros and cons of each presentation style. We also asked them for their thoughts on how they may use such overviews to ask for further feedback. The interviews were conducted in person when possible. When a participant requested a remote session, it was carried out over Zoom where the researcher shared the probes through screen-share and audio-recorded the interview. Each session lasted approximately an hour and all participants were offered CAD 20 gift cards.

**Participants:** We recruited 12 new participants, P09 to P20, (8F/4M) through personal contacts of the researchers and through word-of-mouth. Our participants' ages ranged from 18-44, and they were a mix of university students (undergraduate and graduate) as well as professionals from the software industry, engineering, and art with several years of experience. The 4 students came from the CS and Biology departments. Each participant had experience learning technical skills informally during their careers.

**Data Analysis:** The interviews were audio recorded with the participants' consent and later transcribed. Two researchers analyzed the qualitative data from the interviews using an inductive analysis approach, similar to the first study. We began with open coding [63] as transcripts became available and iterated with the team to arrive at a final coding scheme. While coding, we considered participants' perceptions about more nuanced aspects related to reflections and planning, using temporal, topic-based and resource-based presentations in the three overviews. Following this, we performed axial coding and diagramming to explore themes around our research question. The key insights are synthesized into 3 main themes around how learners perceived the interactive overviews.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2024.3366469

### 4.6 Learners' Perceptions of the Interactive Overviews

Overall, participants expressed that the Temporal Overview and Resource-based Overview were "*visually pleasing*" and easier to interpret. Participants usually began by scanning for the overarching topics, followed by sub-topics. Next, they focused their attention on their own knowledge and areas of weakness or pending action items. Finally, participants desired the ability to edit the overviews to make them more suitable for seeking or offering feedback.

#### 4.6.1 Useful to evaluate the quality of time spent learning

Participants considered the weekly overviews to be helpful in serving as a reminder of the recent learning activities. Most participants (8/12) ranked the Temporal Overview (see Fig. 2) as their first preference as it gave them a quick view of recent efforts, and they could look up concepts within resources using date as a cue without necessarily having to recall terminologies. Participants expressed that they could use granular time-based information to consider how well they are able to pursue different goals within their day, week, or month. P10 shared how time-based overview could help self-learners plan their time better, especially if it offered them the flexibility to change the duration of the overviews: *"If I were to learn something over a month, I would first create weekly to-dos. After completing 4 weeks and before starting a new month, I would need a monthly picture [of things I have done]."*

Similarly, P15 added that a time-based overview over a larger span, such as several months, could also help assess resources they *"have been visiting, but switched up later"*, and they could re-evaluate and reconsider their choices. P13 further shared how the Temporal Overview could reveal how effectively one engages with a resource. For example, *"visiting YouTube to watch one video but spending next half-hour watching cat videos"* would show up as a gap in learning and wasted time. P13 added that *"there isn't necessarily a correlation between time and what I learned"* without also considering, *"did I encounter a ton of problems and do I have to adjust my expectation when I can finish it?"*

Five other participants also voiced similar concerns about having a time-based overview as it could downplay the activities undertaken during a specific period and lead to ambiguity in interpretation (e.g., was the resource helpful or was it difficult?). Additionally, the granular hourly information could also become tedious to analyze for those who only wished to see a cumulative account of time (e.g., total time spent on a topic). Four participants mentioned that the time-based overview combined with or used in a sequence with resource-based overviews could better cater to their needs.

#### 4.6.2 Helpful to reflect on resources and evolving objectives

Among the three probes, two were designed to highlight different aspects of resources used: while the Resource-based Overview showed a high-level classification of the resource type, the Topic-based Overview highlighted keywords and concepts from within collections of resources. We found that 5/12 participants ranked Resource-based as their first choice for overviews, while 7 others ranked it as their second preference, next to the Temporal Overview.

Half of the participants (6/12) preferred to first identify broad topics that they had learned about, followed by sub-categories. As the groups were distinct and labeled in the Resource-based Overview, participants found it easier to understand and interpret, and helpful for identifying the broad categories: *"I feel like it is important for me to know generally what topics I have been looking into, and their subtopics possibly." P11*. Additionally, participants' comments revealed that they wanted to have control over the topics appearing in the overviews. They expressed reluctance over automatically generated sub-topics in Topic-based Overview, stressing that an indicator of the fraction of content they had consumed or found relevant would be more useful. A summary of all the content within visited webpages, which they may not have read entirely or found entirely helpful, would be less desirable. P18 said that learners needed to be more in control of identifying topics and their relevance, adding *"how [the topics] show up can be automated, but I want to decide what [content] is being shown [in the overview]."*

The majority of participants (9/12) wanted to focus their attention on specific parts of the overview to identify areas of weakness. Our use of the color red to represent items that required attention was particularly helpful for participants. As an improvement, five participants shared that *"it would be great to have a way to see outstanding questions"* (P13), short notes to guide their attention to resources that need to be revisited. Participants added that resources that deal with the same category or subcategory of themes needed to be distinguishable from one another, such as through the use of indicators of the required action (e.g., to read, to code, etc). Along the same lines, seven participants desired to see a better connection between learning objectives and resources to better plan their next session. In the context of project-driven learning, participants expressed a desire to see the relevance of resources concerning the problems they were pursuing: *"I would look for a relation between the to-do list and what I have done [based on] the resources [in the overview]. It would be helpful to check if I am missing any items from my list."(P16)*

Although all of our participants found the idea of an integrated to-do list to be helpful, they said that feedback from mentors or peers would give them the most clarity on their learning approach. Next, we describe how the participants speculated on using the different learning overviews to obtain feedback.

#### 4.6.3 Desire to share learning histories for feedback

As previous research has shown, when learners rely on others to ask for help, they can struggle with describing their question or may not be able to articulate it using an accurate vocabulary [24]. We asked participants if they would willingly share the learning history overviews with others to seek feedback and the kind of information they would share or hide. All participants expressed that they would be strongly willing to share their learning histories with a more knowledgeable peer or mentor that they trusted and who could offer advice on their learning approach. They wanted validation of the resources they had selected and be able to *"write down a question, then move on to learning something else, then come back to it or ask a colleague or an expert"* (P13).

Most participants (9/12) described different ways in which they would alter the overviews to make them more suitable for seeking feedback. For example, P09 shared how she would add annotations to the overview: *"I want to be able to send this [overview] to my peer and have them check out my bookmarks - I might add a note to let them know how the resource helped me."* P10 and P12 both mentioned that in technical subject areas, they would also add references to their implementation attempts or any example questions they may have resolved: *"[For an applied skill] I would like to see a separate category [in Resource-based Overview] for practice sessions [arranged] by topics."* P18 added that while implementation could be an important detail to add for useful feedback, learners should have a choice to share it optionally on request.

Some others (P16, P19, P20) also mentioned removing extraneous items from their learning histories. P20 shared how she was *"not comfortable sharing how [she] learned something, but if there was a way to filter out specific details [...] then that would be good."*. She preferred to abstract the information down to essentials, such as themes, titles, links, and sequence of visits. Overall, participants considered overviews to be helpful for *"show and tell"* and wanted to have editing controls to make them more suitable for feedback.

## 5 Study 3: Experts' Insights for the Design and Deployment of Self-Monitoring Tools in the Wild

While our probe-based studies provided valuable insights into learners' self-monitoring needs, we wanted to ensure that we are capturing a more holistic and pragmatic perspective for developing self-monitoring tools for learning. We sought insights from HCI and information visualization experts who had direct experience in deploying and studying various self-monitoring tools in real-world settings. These experts, well-versed in the nuances and challenges of actual tool implementation, provided invaluable feedback that extended beyond the scope of learner insights alone. Their perspectives were instrumental in validating and refining our initial findings, ensuring that the design of the

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2024.3366469

self-monitoring tool aligns with real-world application and usability. This collaboration helped in identifying key design requirements and potential barriers that might not be evident in controlled study environments. It also allowed us to consider broader implications, such as scalability, adaptability, and long-term effectiveness of the tool in diverse learning contexts.

### 5.1 Study Procedure and Participants

We recruited 6 experts in the field of HCI and Information Visualization (4F/2M) through personal contacts and by snowball sampling. Each expert was a researcher affiliated with a different university and had experience designing self-monitoring technologies in domains such as health, well-being, and productivity. The experts possessed a minimum of five years of research experience in their respective areas of focus, with extensive involvement in designing and conducting various self-monitoring user studies.

The self-monitoring projects described by experts included various aspects, such as managing stress and productivity during knowledge work, exploring data collection techniques, enhancing the accessibility of monitoring tools, and collaborating with healthcare providers to monitor and manage health conditions. The projects deployed a wide spectrum of interventions, including developing narrative techniques that involve challenging assumptions through the application of questioning methods derived from Cognitive Behavioral Therapy frameworks and incorporated various technological interventions across multiple devices, ranging from sensor-based tools to mobile applications, web interfaces, and virtual reality experiences. Additionally, the experts had hands-on experience conducting field studies involving functional self-monitoring tools for personal use (3/6) and overseeing studies using citizen healthcare data (2/6). Furthermore, our panel of experts brought valuable insights from their interactions with diverse participant groups, including students and programmers, individuals seeking healthcare services, such as those in addiction recovery, stroke survivors, older adults, and marginalized populations.

Inspired by the idea of data triangulation and using multiple perspectives to add rigor to our qualitative analysis, we designed this study to gather insights from experts on designing self-monitoring tools [46, 53]. We conducted semi-structured interviews remotely over Zoom. Each interview was audio-recorded and lasted approximately 30-40 minutes. During these interviews, we shared our findings from the user studies discussed, and we asked each interviewee to reflect on their own experiences with self-monitoring projects. We encouraged them to highlight the most pertinent factors that they believed would contribute to our ongoing efforts in eliciting design requirements for a self-monitoring tool tailored to informal learners. In particular, we asked them for advice on building and deploying a self-monitoring tool based on our current requirements. Subsequently, we transcribed, coded, and analyzed the interview data using an inductive analysis method [63].

In the following sections, we share our findings regarding three main aspects: first, assessing the adequate amount of data to gather and strategies to ease the data collection process; second, exploring users' viewpoints on self-monitoring and strategies to enhance their monitoring experience. Lastly, we delve into the topic of establishing trust with users during the field deployment of self-monitoring tools and encouraging their active engagement with the interventions.

### 5.2 Considerations for Data Collection Techniques

The majority of experts discussed research projects that necessitated participants to manually record their personal data. Common tasks for the study participants included documenting their emotions and stress levels, responding to inquiries about their dietary habits, or providing self-assessments of their productivity. To support participants in these logging activities, notifications or reminders were frequently provided, and participants were granted access to visual representations of their recorded data for reflections (further details on this are provided in the subsequent subsections).

We posed a question regarding data collection methods, prompting the experts to describe both effective and ineffective approaches they encountered in their studies. Two recurring themes emerged: 1) determining the appropriate amount of data to gather and 2) devising methods to alleviate the burden associated with data collection.

#### 5.2.1 How much data is sufficient for generating insights?

To help participants extract meaningful insights from the personal data collected during the studies, the experts (3/6) mentioned their reliance on questionnaires commonly employed in behavioral interventions. While increasing the number of questions posed to participants could enhance the data's utility for future reflections, experts acknowledged that this approach could become burdensome for participants. As E5 pointed out, even in a simple self-monitoring scenario like food journaling, manual logging could become overwhelming as *"every time you make a decision about eating something, you have to think about "why I choose this food when I made the decision. How do I decide this is a time to eat and how much to eat?" It's actually a lot."*

Furthermore, E3 observed that the questionnaires used for research purposes, as opposed to interventions, sometimes influenced how users engaged in self-monitoring. However, users didn't always perceive these questionnaires as personally valuable. This highlights a tension between data collected for personal use and data collected for research. Researchers may need to ask questions to gather contextual information, such as when it's suitable for participants to log data and how they feel about the process, and tailor their research or design interventions accordingly. For example, as E3 described, *"it's tricky to isolate whether or not it's the tool's problem or if it's just the wrong time for the participants"*. In E1's experience, it was usually sufficient to ask participants up to three carefully chosen questions that they could answer comfortably and found useful for subsequent reflections.

Experts emphasized the importance of data logging and having the flexibility to record personally meaningful information. In one of E4's projects, which was focused on collecting citizen data about limitations in social life and functional abilities arising from diseases, E4 noted that factors like technological literacy, spoken language, and the extent of pain's impact on participants' lives influenced their engagement in data contribution and utilization. Users also needed a clear understanding of how *"contributing this information will help other people like [themselves] and [being] able to see other people's information that will help [them] in return."*. However, cultivating this understanding of data collection's benefits could be challenging and require targeted efforts to convey. The experts noted that once users recognized the value of data collection, ongoing support was necessary to maintain their involvement in these activities. We discuss examples of possible ways of extending support in the following subsection.

#### 5.2.2 Strategies to alleviate the burden of data collection

The experts participating in our study emphasized the importance of establishing the timing for data logging and incorporating moments for reflection, as these aspects are often not at the forefront of people's attention. However, they also noted that reminders, prompts, or notifications could be counterproductive if poorly timed or overly frequent. As a potential solution, several experts (E1, E2, E3) recommended granting users the flexibility to decide when to engage in self-monitoring activities. Apart from allowing users to customize the timing of reminders, the experts emphasized the importance of identifying and designing for the collection of personally relevant data, such that manual data collection becomes more intrinsically motivating. E2 used food tracking as an example: *"the way that people track foods, it can be all very different. Some people focus on breakfast tracking, some people focus on a restaurant that they visited, and others focus on a particular food that they are interested in and so on.."*

However, there are some important considerations to take into account. E2 noted that individuals who derive the most benefit from personalization features are typically *"power users"* or experts, while newcomers may require time to learn and adapt to these personalized features. Moreover, the push for more structured approaches may not accommodate the inherent uncertainties present in certain situations. E6 pointed out that in one of her studies, a participant was dealing with addiction, making the circumstances inherently uncertain and challenging for them to predict or identify what to log in the moment. In such cases, participants engaged in self-experimentation under the guidance of clinicians to discover effective strategies, given the unpredictable nature of their situations.

In the context of simplifying data collection, three experts in our study explored different input modes to facilitate the process. They noted that combining various modalities (such as speech, touch, and graphic indicators) could help overcome challenges in capturing data. E5, for example, explained that in her work speech input proved valuable for collecting rich contextual data compared to text input. However, speech input appeared most suitable for private and quiet environments and straightforward interactions. As such, an optimal approach to leverage input methods would be to use them in combination (e.g., speech and touch) to facilitate natural and productive interactions.

## 5.3 Considerations when Designing for Reflections

In our discussions with the experts, we asked about designing for reflections, focusing on factors that they personally found significant when creating self-monitoring tools that facilitate reflection. In our definition, we characterized reflections as the process through which individuals can analyze past events, gain insights, and identify areas for potential change. Two recurring themes emerged from the responses. First, the experts noted that participants in certain field studies developed a negative connection with personal data, especially related to their ongoing personal challenges, and acknowledged the difficulty in designing for reflection during these moments. The second theme revolved around approaches to address this issue, with experts presenting various strategies. These ideas aimed to involve users and lead them through iterative goal-setting and fruitful self-experimentation.

### 5.3.1 Self-perceptions affecting the accuracy of personal data

According to E3, users often experience self-criticism when reviewing personal data related to productivity or time management. This highlights the importance of helping users disassociate their data from their self-worth, ensuring that data fosters awareness without triggering self-judgment. E6 referred to this phenomenon as *"self-rejection"* and elaborated that participants in her study often anticipated negative records before viewing their data and were sometimes surprised to see a greater incidence of evidence associated with positive outcomes. E6 also noted that some users would backlog their entries or leave gaps. She noted that users may feel reluctant to log when they were feeling negative or when there was negative data and preferred to *"wait for a happy feeling so that [they] could log it."* . E3 further added that during manual data tracking, users might initially log more negative events and later alter their behavior to record more positive events, effectively deceiving themselves, a behavior he referred to as *"productivity theatre"* [7]. This alteration often occurred when users believed their data would be reviewed by someone else, such as their managers, and it influenced their logging behavior to appear more productive, defeating the purpose of self-monitoring for self-improvement.

### 5.3.2 Supporting reflections using guided iterative goal-setting

Among the methods the experts recommended for facilitating self-reflection, the idea of employing structured reflective queries and step-by-step goal-setting assistance was the most popular. To address situations where certain individuals may possess negative self-perceptions during self-reflection, E6 introduced a concept based on *"self-experimentation"*, which prioritizes the exploration of progressive positive changes. E6 applied this approach in the context of learning, where she asked her students, who came from diverse disciplines and learning backgrounds, to reflect on how they had improved. By prompting them to compare their past performance with their current state over the course of a term, she aimed to shift their mindset toward recognizing their achievements. This allowed students to assess how their learning strategies had contributed to their improvement, strengthening their emotional state and enabling more strategic thinking about future improvements. E3 suggested borrowing from other frameworks, such as Socratic questioning commonly used in CB to *"challenge assumptions or to broaden people's perspective... like if a friend were to give you some advice about the situation, like an outside perspective"*

P4 discussed one of her workshop experiences where she saw how individuals could develop a negative outlook toward their personal data. She proposed integrating the two approaches of comparing with peers at similar levels and engaging in reflective questioning. Participants would not compare their performance data with others but instead compare their reflective responses to the question: *"What are the three most important things I learned from this course?"* By viewing other participants' responses, individuals could gain insights into important aspects and significantly shift their perspectives.

Additionally, both E1 and E6 stressed the importance of iterative goals setting. They emphasized that users benefit from the exercise of trying to break down larger objectives into smaller tasks following the "SMART goals" framework where participants first formulate a *"big smart goal and then [...] fill out like 3 or 4 mini-steps that they could do to reach that smart goal."*. However, users may struggle to conceptualize these smaller steps, necessitating examples and mentor support to assist in this process. For instance, users may have difficulty breaking down a goal like "going running" into smaller, manageable steps. Providing guidance on setting achievable mini-goals, such as preparing running shoes or getting dressed for a run, can make the process more attainable and effective.

## 5.4 Considerations for Field Studies

We sought advice from the experts in our study about factors to consider when deploying a functional self-monitoring tool for field studies and precautions that should be taken. The predominant themes revolved around: 1) building trust and safeguarding the participants' welfare throughout the data collection process; and 2) enhancing engagement with self-monitoring interventions.

### 5.4.1 Building trust and safeguarding participant's welfare

A primary factor highlighted by all experts centered on ensuring that study participants and potential users fully comprehend the implications of their data consent. They stressed the importance of assuring users about data safety by detailing the *"the purpose, ask and benefits"* of collecting their self-monitored data for the study. Three experts highlighted participants being unable to *"realize how personal or how sensitive those data could be [...] And what are some of the ramifications of sharing those data with you, the researcher, long term."* E5 elucidated this problem with an example from one of her field studies on food journaling, where her participants initially had no hesitation to log their eating habits, but later realized that they were, *"disclos[ing] too much personal information in [their] food decision* by revealing information about their *"personal life, work, partner, relationship and concerns about body image"*.

All experts emphasized taking measures to establish trust with the participants by allowing them enough time and *"privacy to go over some of the important content"* (E2). Experts also advised carrying out one-on-one interactions to build trust, as well as offering examples and snapshots of automatically collected data. Additionally, in cases where data needed to be shared with others as part of the study, E6 suggested using the concept of *"trust circles"*. This involved categorizing individuals into circles based on their level of trustworthiness, with the innermost circle representing those most trusted with the data and those farther from the center indicating decreasing levels of trust.

### 5.4.2 Enhancing engagement with self-monitoring interventions

Experts raised additional considerations to take into account before deploying a self-monitoring tool for field studies. When asked about the accuracy and bias in the collected data, the experts warned that certain data collection techniques were more error-prone. For example, the experts shared numerous instances where data manually entered by participants was inherently biased and less accurate.

Although the consensus among the experts was that participants benefit from assistance through reminders and prompts for self-monitoring,

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2024.3366469

E3 cautioned about participants' inherent motivation and genuine engagement with these interventions. E3 pointed out that these techniques could alter the natural engagement with the tool, making it challenging to discern if the engagement is driven by genuine value. In another example, E3 elaborated that if engagement strategies involve sharing data with others, participants may be less candid due to fear of judgment, such as knowledge workers sharing productivity data with their managers. Based on his research, E3 noted that different parties involved may have varying definitions of productivity and standards, causing knowledge workers to align with their manager's expectations, giving the illusion of higher productivity, and ultimately undermining the purpose of self-monitoring.

In another example of inaccuracies in manually collected data, E6 mentioned a participant who would backlog her data on specific days, disrupting the timeline visualization as the software couldn't account for this backlog. Furthermore, in the context of learning, E6 discussed how *"learning is so difficult to quantify"* as the details needed to assess progress were based on the quality of the reflection entries that students submitted as a form of self-assessment based on the artifacts they produced over the course of a term.

## 6 Discussion

We have taken a learner-centered approach [62] to understand what matters to informal learners when capturing and visualizing self-monitoring data and reflecting on learners' needs and experiences. We explored several self-monitoring dimensions and visual overviews through paper-based mockups, followed by interactive probes inspired by productivity, health, and wellness tracking tools. [15]. Our probe-based studies [67] in the lab provided valuable insights into learners' self-monitoring perceptions and needs. Furthermore, our follow-up interviews with experts offered alternative interpretations, revealing important design tensions within our findings and expanding our understanding of designing and deploying self-monitoring tools. Below, we reflect on the key insights and the design tensions that emerged and formulate design requirements for self-monitoring interventions for informal learners.

### 6.1 Design Tensions in Data Capture and Visualization Techniques for Self-Monitoring Learning

Our three-part study, exploring learners' perceptions (sections 3 and 4) and collecting insights from experts (section 5), exposed divergent viewpoints on the challenges of designing visualizations for self-monitoring in learning contexts. First, to gain insights and feedback from trusted others, learners expressed a desire for customized overviews and selective data sharing, emphasizing control over essential information (section 4.6.3). In contrast, experts recommended extensive behavioral questionnaires to foster awareness and insight generation through effective reflection practices, albeit acknowledging their long-term sustainability issues. In essence, **this creates a design tension between enabling learner autonomy in data sharing and recording and designing interventions that optimize opportunities for reflection**. Editable visual overviews could empower learners by offering control over sharing and refining data and could facilitate reflective conversations with a trusted other.

Second, we learned that individuals may hold negative self-perceptions about the behavior they are self-monitoring, potentially influencing or biasing data collection and interpretation. While semi-automatic systems promise consistency, the aspect of manual recording introduces concerns about compromised accuracy, prompting the question of **whether to design for accuracy or support positive self-perceptions to encourage engagement with self-monitoring, even if accuracy is compromised**. Prior work, especially in contexts related to learning, suggests that consistency may outweigh the benefits of accuracy in achieving behavior change [8]. In this regard, semi-automated data collection could enhance a learner's engagement and consistency in self-monitoring behaviors by establishing a foundational baseline through automatic recording and empowering users to choose what and how to log for generating valuable insights through manual approaches. Additionally, fostering participants' understanding of the purpose and significance of collecting personal data may encourage them to participate persistently.

Finally, learners conveyed that they typically do not incorporate goal-setting into their routine, while experts advocated for guided iterative goal-setting to encourage self-reflection. Experts proposed structured approaches such as "SMART goals" (section 5.3.2), while learners expressed a preference for discovering and revisiting topics and subtopics and making associations between resources and tasks through a retrospective exploration of visual overviews - thus revealing **the tension between balancing the experts' structured goal-setting approach with the learners' preference for flexibility and adaptability in the learning process**. In this regard, interactive visualizations could provide useful insight for informal learners and enable them to engage in self-reflection on various aspects of learning, such as evaluation of the time utilized for different goal pursuits or reconsideration of choices made in terms of resources and learning strategies.

### 6.2 Key Requirements for Designing Self-Monitoring Interventions for Informal Learning

Keeping in mind that technology design involves inherent trade-offs, we now present recommendations that aim to strike a balance between competing goals, as revealed through the design tensions from the probe-based elicitation and expert validation studies.

#### 6.2.1 Support automatic tracking and interactive overviews

The visual overviews in our study were considered helpful because they could serve as a quick at-a-glance summary [19] of learning efforts undertaken over a period through minimal learner intervention. However, to make visual overviews useful, any labeling achieved through automation should communicate how the labels were generated. Additionally, the intervention should offer transparency regarding the data collected, allowing learners to decide the extent of data collection and its timing. Moreover, visualizations could be made more effective in drawing and guiding learners' attention by showing data in categories, such as broad topics or relevant resource meta-data, that learners can recognize with ease. Learners should be supported in obtaining different insights through manipulation and interaction (e.g., select, filter, zoom) with the presented data [19, 38, 71].

#### 6.2.2 Provide manual markers and personally relevant filters

To make the overviews personally relevant, learners should be offered semi-automatic support, such as through automatically suggested tags, along with the flexibility to specify their own tags or mark up and filter the resources according to metrics such as perceived usefulness, level of difficulty, and relevance to goals. Learners who can identify nuances between resources should be allowed to indicate the variations. For example, resource-specific indicators could help learners decide which resources to revisit or show the extent to which learners have consumed a resource (or where they have left off), outstanding questions, or action items (e.g., "to read", "to code", "to watch") associated with specific resources. Given the challenge of assessing the quality of a resource at the outset, learners should have the ability to account for this uncertainty. This can be achieved by facilitating easy corrections, enabling edits to judgments about resource utility, or revealing iterations in the tracked metric (e.g., perceived difficulty or open questions) as they evolve over time.

#### 6.2.3 Reserve goal tracking for advanced stages of learning

While goal-tracking and checking for goal alignment was unanimously considered a useful activity by all the participants in our study, there were apprehensions that such interventions may not be feasible for beginners. Learners may begin with broad goals, which may become more specific over time. Goal-tracking, therefore, should be adaptive and suited to the stage of learning [42]. For example, learners could be given access to simple goal trackers and reminders in the early stages of their learning. Learners could also be supported in shaping their goals, either by breaking down larger goals into smaller ones or by piecing together smaller foreseeable goals to determine an overall direction of pursuit over time. More sophisticated goal-tracking, such as fine-tuned task identification, should be reserved for more advanced learners who may be at the application or implementation stage of their learning.

#### 6.2.4 Allow tailoring of visual overviews for feedback

Self-monitoring can be made more effective with occasional feedback from experts or knowledgeable peers [34]. We learned that visual overviews can be used to solicit feedback on learning strategies as they reveal the learner's pathways and attempts. However, our participants wished to tailor their overviews based on the type of feedback they wanted. This suggests that learners could benefit from the flexibility in determining the extent of detail to share for feedback. Interactive overviews should include ways to allow learners to add more context with further annotations or remove details they consider redundant or irrelevant to the desired feedback. Furthermore, considering that self-monitored data may include personal information, it should be feasible to customize the level of data sharing according to the position of the feedback provider within the learner's circles of trust.

#### 6.2.5 Limitations and Future work

Although we recruited participants from different backgrounds and professions, future works should consider a more diverse set of learners of computational skills who may have different learning styles. Since the context we used in our scenarios was limited to online mediums only and our questions were focused on individuals learning technical skills, whether our results will generalize to other learning scenarios, such as in formal teacher-led settings or collaborative settings, should be further investigated. Moreover, the data we showed in the probes were curated by the researchers. More qualitative studies are required to uncover the nuances of self-monitoring in domain-specific informal learning through the use of domain-specific content and ethical ways of using participants' own data. Future studies could also expand the informal learning context to physical resources and artifacts and explore how digital methods could augment physical tracking. Future studies could use experimental methods, observational methods, or in-situ data collection methods, such as experience sampling, design probes, or journaling, to triangulate the responses.

## 7 Conclusion

Our iterative design approach has contributed insights into informal learners' perceptions of self-monitoring and revealed that learners in our study prefer automatically generated interactive visual overviews of learning activities for their ability to enhance awareness of learning processes. Learners expressed a willingness to actively participate in data collection to make it more suitable for reflections and planning. Furthermore, the validation by experts confirmed our findings, revealed design tensions, and provided additional insight into factors to consider when designing and developing self-monitoring tools for deployment. Our work opens up several opportunities for future research and advocates for employing learner-centered approaches to understand and cater to the needs of informal and self-directed learners.

## 8 Acknowledgments

We thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding this research.

## References

[1] P. Abtahi, V. Ding, A. C. Yang, T. Bruzzese, A. B. Romanos, E. L. Murnane, S. Follmer, and J. A. Landay. Understanding Physical Practices and the Role of Technology in Manual Self-Tracking. *Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(4):1–24, Dec. 2020. 2, 3, 4

[2] A. Agrahri. Index of Best AI/Machine Learning Resources. *https://medium.com/hackernoon/index-of-best-ai-machine-learning-resources-71ba0c73e34d*, May 2019. 1

[3] J. Ahn, F. Campos, M. Hays, and D. DiGiacomo. Designing in Context: Reaching beyond Usability in Learning Analytics Dashboard Design. *Journal of Learning Analytics*, 6(2):70–85, 2019. 3

[4] A. Ayobi, P. Marshall, and A. L. Cox. Reflections on 5 Years of Personal Informatics: Rising Concerns and Emerging Directions. In *CHI Conference on Human Factors in Computing Systems*, pp. 2774–2781, 2016. 4

[5] A. Ayobi, T. Sonne, P. Marshall, and A. L. Cox. Flexible and mindful self-tracking: Design implications from paper bullet journals. In *CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2018. doi: 10.1145/3173574.3173602 4

[6] A. Bakharia, K. Kitto, A. Pardo, D. Gašević, and S. Dawson. Recipe for success: lessons learnt from using xAPI within the connected learning analytics toolkit. In *ACM LAK*, pp. 378–382, 2016. 3

[7] S. Bhaimiya. Workers say they're engaging in "productivity theater" for over 10 hours a week to appear busy, a survey shows, May 2023. 9

[8] A. Bruhn and H. P. Wills. Emerging research and development in technology-based self-monitoring. *Emerging research and issues in behavioral disabilities*, 30:51–68, 2018. 10

[9] C. J. Cai and P. J. Guo. Software developers learning machine learning: Motivations, hurdles, and desires. In *VL/HCC*, pp. 25–34. 1

[10] E. A. Canning and J. M. Harackiewicz. Utility value and intervention framing. *The Cambridge handbook of motivation and learning (pp. 645–662)*, 2019. doi: 10.1017/9781316823279.027 4

[11] R. Chaudhury, P. J. Guo, and P. K. Chilana. "There's no way to keep up!": Diverse Motivations and Challenges Faced by Informal Learners of ML. In *VL/HCC*, pp. 1–11, 2022. doi: 10.1109/VL/HCC53370.2022.9833100 1, 2

[12] L. Chen, M. Lu, Y. Goda, A. Shimada, and M. Yamada. Learning analytics dashboard supporting metacognition. *Springer International Publishing*, pp. 129–149, 2021. doi: 10.1007/978-3-030-65657-7_8 2, 3

[13] P. K. Chilana, R. Singh, and P. J. Guo. Understanding Conversational Programmers: A Perspective from the Software Industry. In *CHI Conference on Human Factors in Computing Systems*, pp. 1462–1472. ACM. 1, 2

[14] J. Cho, T. Xu, A. Zimmermann-Niefield, and S. Voida. Reflection in Theory and Reflection in Practice: An Exploration of the Gaps in Reflection Support among Personal Informatics Apps. In *CHI Conference on Human Factors in Computing Systems*, pp. 1–23, 2022. 2, 3

[15] E. K. Choe, S. Abdullah, M. Rabbi, E. Thomaz, D. A. Epstein, F. Cordeiro, M. Kay, G. D. Abowd, T. Choudhury, J. Fogarty, B. Lee, M. Matthews, and J. A. Kientz. Semi-Automated Tracking: A Balanced Approach for Self-Monitoring Applications. *IEEE Pervasive Computing*, 16(1):74–84, Jan. 2017. doi: 10.1109/MPRV.2017.18 2, 3, 4, 10

[16] E. K. Choe, B. Lee, H. Zhu, N. H. Riche, and D. Baur. Understanding Self-reflection: How People Reflect on Personal Data through Visual Data Exploration. In *Pervasive Computing Technologies for Healthcare*, pp. 173–182, 2017. 3

[17] D. S. K. Chris Elsden and A. C. Durrant. A Quantified Past: Toward Design for Remembering with Personal Informatics. *Human–Computer Interaction 2016*, pp. 518–557. doi: 10.1080/07370024.2015.1093422 3

[18] B. Dorn and M. Guzdial. Learning on the job: Characterizing the programming knowledge and learning strategies of web designers. In *CHI Conference on Human Factors in Computing Systems*, p. 703, 2010. 2

[19] D. A. Epstein, T. Blascheck, S. Carpendale, R. Dachselt, and J. Vermeulen. Challenges in everyday use of mobile visualizations. In *Mobile Data Visualization*, pp. 209–240. Chapman and Hall/CRC, 2021. 4, 5, 10

[20] D. A. Epstein, A. Ping, J. Fogarty, and S. A. Munson. A lived informatics model of personal informatics. In *UbiComp*, pp. 731–742, 2015. 2, 3

[21] K. Falkner, R. Vivian, and N. J. Falkner. Identifying Computer Science Self-regulated Learning Strategies. In *ITiCSE '14*, pp. 291–296. 2

[22] A. Fekete, J. Kay, J. Kingston, and K. Wimalaratne. Supporting Reflection in Introductory Computer Science. *SIGCSE Bulletin 2000*, pp. 144–148. 2

[23] C. Feustel, S. Aggarwal, B. Lee, and L. Wilcox. People Like Me: Designing for Reflection on Aggregate Cohort Data in Personal Informatics Systems. *Interactive, Mobile, Wearable and Ubiquitous Technologies*. 3

[24] G. W. Furnas, T. K. Landauer, L. M. Gomez, and S. T. Dumais. The vocabulary problem in human-system communication. *Communications of the ACM*, 30(11):964–971, 1987. 3, 7

[25] P. J. Guo. Older Adults Learning Computer Programming: Motivations, Frustrations, and Design Opportunities. In *CHI Conference on Human Factors in Computing Systems*, pp. 7070–7083. 2

[26] J. D. Hansen and J. Reich. Democratizing education? Examining access and usage patterns in Massive Open Online Courses. *Science*, 350(6265):1245–1248, 2015. 2, 3

[27] T. Hogan, U. Hinrichs, and E. Hornecker. The Elicitation Interview Technique: Capturing People's Experiences of Data Representations. *IEEE Transactions on Visualization and Computer Graphics*, pp. 2579–2593, 2016. doi: 10.1109/TVCG.2015.2511718 3

[28] D. Huang, M. Tory, B. Adriel Aseniero, L. Bartram, S. Bateman, S. Carpendale, A. Tang, and R. Woodbury. Personal Visualization and Personal

Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 21(3):420–433, 2015. doi: 10.1109/TVCG.2014.2359887 2, 3

[29] Z. Huang, Q. He, K. Maher, X. Deng, Y.-K. Lai, C. Ma, S.-F. Qin, Y.-J. Liu, and H. Wang. Speechmirror: A multimodal visual analytics system for personalized reflection of online public speaking effectiveness. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):606–616, 2024. doi: 10.1109/TVCG.2023.3326932 3

[30] J. Jones, Y. Yuan, and S. Yarosh. Be consistent, work the program, be present every day: Exploring technologies for self-tracking in early recovery. *Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(4):1–26, 2021. 4

[31] J. A. Jones. Scaffolding self-regulated learning through student-generated quizzes. *Active Learning in Higher Education*, 20(2):115–126, 2019. 3

[32] J. Kay, K. Bartimote, K. Kitto, B. Kummerfeld, D. Liu, and P. Reimann. Enhancing learning by open learner model (olm) driven data design. *Computers and Education: Artificial Intelligence*, 3:100069, 2022. 3

[33] C. Kelly, B. Lee, and L. Wilcox. Self-tracking for Mental Wellness: Understanding Expert Perspectives and Student Experiences. *CHI Conference on Human Factors in Computing Systems*, Feb. 2017. 3

[34] K. Kiani, P. K. Chilana, A. Bunt, T. Grossman, and G. Fitzmaurice. “I Would Just Ask Someone”: Learning Feature-Rich Design Software in the Modern Workplace. In *VL/HCC*, pp. 1–10, 2020. 11

[35] Y.-H. Kim, E. K. Choe, B. Lee, and J. Seo. Understanding Personal Productivity: How Knowledge Workers Define, Evaluate, and Reflect on Their Productivity. *CHI Conference on Human Factors in Computing Systems*, 2019. 3

[36] Y.-H. Kim, J. H. Jeon, E. K. Choe, B. Lee, K. Kim, and J. Seo. TimeAware: Leveraging Framing Effects to Enhance Personal Productivity. *CHI Conference on Human Factors in Computing Systems*, 2016. 2

[37] Y.-H. Kim, J. H. Jeon, B. Lee, E. K. Choe, and J. Seo. OmniTrack: A Flexible Self-Tracking Approach Leveraging Semi-Automated Tracking. *Interactive, Mobile, Wearable and Ubiquitous Technologies*, July 2017. 2, 3

[38] Y.-H. Kim, B. Lee, A. Srinivasan, and E. K. Choe. Data@ hand: Fostering visual exploration of personal data on smartphones leveraging speech and touch interaction. In *CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2021. doi: 10.1145/3411764.3445421 3, 10

[39] M. S. Knowles. Self-directed learning: a guide for learners and teachers. *ERIC*, 1975. 1, 2

[40] A. J. Ko, B. A. Myers, and H. H. Aung. Six Learning Barriers in End-User Programming Systems. In *VL/HCC 2004*, pp. 199–206, Sept. 2004. 2

[41] R. Langner, M. Satkowski, W. Büschel, and R. Dachselt. Marvis: Combining mobile devices and augmented reality for visual data analysis. In *CHI Conference on Human Factors in Computing Systems*, CHI ’21. New York, NY, USA, 2021. doi: 10.1145/3411764.3445593 3

[42] D. Loksa, L. Margulieux, B. A. Becker, M. Craig, P. Denny, R. Pettit, and J. Prather. Metacognition and Self-Regulation in Programming Education: Theories and Exemplars of Use. *ACM Transactions on Computing Education*, p. 3487050, Feb. 2022. doi: 10.1145/3487050 2, 10

[43] G. Lupi and S. Posavec. *Dear Data*. Chronicle books, 2016. 3

[44] L. A. Lyon, C. Clayton, and E. Green. Tinkering in the Wild: What Leads to Success for Female End-User Programmers? In *VL/HCC*, pp. 125–129, Oct. 2018. doi: 10.1109/VLHCC.2018.8506561 2

[45] M. C. Maphalala, R. G. Mkhasibe, and D. W. Mncube. Online learning as a catalyst for self-directed learning in universities during the covid-19 pandemic. *Research in Social Sciences and Technology*, 6(2):233–248, 2021. 1

[46] H. McBride, R. Wiens, M. McDonald, D. Cox, E. Chan, and B. Zumbo. Validity and validation in social, behavioral, and health sciences. In *The Edinburgh Postnatal Depression Scale (EPDS)*, pp. 157–74. 2014. 8

[47] A. N. Meyer, T. Fritz, and T. Zimmermann. Fitbit for developers: self-monitoring at work. *Rethinking Productivity in Software Engineering*, pp. 261–270, 2019. 3

[48] A. N. Meyer, G. C. Murphy, T. Zimmermann, and T. Fritz. Design Recommendations for Self-Monitoring in the Workplace: Studies in Software Development. *ACM Human-Computer Interaction.*, 1(CSCW):1–24. 2

[49] S. Nakazawa and T. Tanaka. Prototype of kanban tool and preliminary evaluation of visualizing method for task assignment. In *International Conference on Control and Automation*, pp. 48–49, 2015. 3

[50] J. Park, H. Lee, S. Park, K.-M. Chung, and U. Lee. Goldentime: Exploring system-driven timeboxing and micro-financial incentives for self-regulated phone use. In *CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2021. 3, 5

[51] E. A. Patall and S. Y. Hooper. The promise and peril of choosing for motivation and learning. *The Cambridge handbook of motivation and learning (pp. 238–262)*, 2019. doi: 10.1017/9781316823279.012 4

[52] M. S. Patel, D. A. Asch, and K. G. Volpp. Wearable devices as facilitators, not drivers, of health behavior change. *Jama*, 313(5):459–460, 2015. 4

[53] M. Q. Patton. Enhancing the quality and credibility of qualitative analysis. *Health Services Research*, 34(5 Pt 2):1189, 1999. 8

[54] R. Pérez-Álvarez, J. Maldonado-Mahauad, and M. Pérez-Sanagustín. Tools to support self-regulated learning in online environments: Literature review. In *European Conference on Technology Enhanced Learning*, pp. 16–30. Springer, 2018. 2, 3

[55] D. Piorkowski, S. D. Fleming, C. Scaffidi, M. Burnett, I. Kwan, A. Z. Henley, J. Macbeth, C. Hill, and A. Horvath. To fix or to learn? How production bias affects developers’ information foraging during debugging. In *ICSME 2015*, pp. 11–20. 2

[56] J. Prather, B. A. Becker, M. Craig, P. Denny, D. Loksa, and L. Margulieux. What do we think we think we are doing? metacognition and self-regulation in programming. In *International Computing Education Research*, pp. 2–13, 2020. 2

[57] I. Radu and B. Schneider. How augmented reality (AR) can help and hinder collaborative learning: A study of ar in electromagnetism education. *IEEE Transactions on Visualization and Computer Graphics*, 29(9):3734–3745, 2023. doi: 10.1109/TVCG.2022.3169980 2

[58] M. B. Rieger and B. Risch. How to maximise spatial presence: Design guidelines for a virtual learning environment for school use. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2517–2526, 2023. doi: 10.1109/TVCG.2023.3247111 2

[59] T. Rohloff, D. Sauer, and C. Meinel. Students’ Achievement of Personalized Learning Objectives in MOOCs. In *ACM Learning @ Scale 2020*, pp. 147–156. doi: 10.1145/3386527.3405918 3

[60] J.-F. Rouet, M. A. Britt, and A. Potocki. Multiple-text comprehension. *The Cambridge handbook of cognition and education*, pp. 356–380, 2019. 4

[61] D. Schott, M. Kunz, T. Wunderling, F. Heinrich, R. Braun-Dullaeus, and C. Hansen. Cardiogenesis4d: Interactive morphological transitions of embryonic heart development in a virtual learning environment. *IEEE Transactions on Visualization and Computer Graphics*, pp. 2615–2625, 2023. doi: 10.1109/TVCG.2023.3247110 2

[62] E. Soloway, M. Guzdial, and K. E. Hay. Learner-centered design: The challenge for HCI in the 21st century. *Interactions*, pp. 36–48, 1994. 2, 3, 10

[63] A. Strauss and J. Corbin. Grounded theory methodology: An overview. *Handbook of qualitative research (pp. 273–285)*, 1994. 4, 6, 8

[64] A. Thudt, D. Baur, S. Huron, and S. Carpendale. Visual mementos: Reflecting memories with personal data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):369–378, 2015. 3

[65] A. Thudt, U. Hinrichs, S. Huron, and S. Carpendale. Self-reflection and personal physicalization construction. In *CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2018. 4

[66] E. Villalobos, I. Hilliger, M. Pérez-Sanagustín, C. González, S. Celis, and J. Broisin. Analyzing learners’ perception of indicators in student-facing analytics: A card sorting approach. In O. Viberg, I. Jivet, P. Muñoz-Merino, M. Perifanou, and T. Papathoma, eds., *Responsive and Sustainable Educational Futures*, pp. 430–445, 2023. 2, 3

[67] J. Wallace, J. McCarthy, P. C. Wright, and P. Olivier. Making design probes work. In *CHI Conference on Human Factors in Computing Systems*, pp. 3441–3450, 2013. doi: 10.1145/2470654.2466473 2, 3, 10

[68] A. Y. Wang, R. Mitts, P. J. Guo, and P. K. Chilana. Mismatch of Expectations: How Modern Learning Resources Fail Conversational Programmers. In *CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2018. doi: 10.1145/3173574.3174085 1, 2

[69] P. H. Winne. Improving Measurements of Self-Regulated Learning. *Educational Psychologist*, 45(4):267–276, Oct. 2010. 3

[70] K. Yatani and K. N. Truong. Bodyscope: a wearable acoustic sensor for activity recognition. In *UbiComp*, pp. 341–350, 2012. 3

[71] J. S. Yi, Y. ah Kang, J. Stasko, and J. A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1224–1231, 2007. 4, 10

[72] H. Zeng, X. Shu, Y. Wang, Y. Wang, L. Zhang, T.-C. Pong, and H. Qu. Emotioncues: Emotion-oriented visual summarization of classroom videos. *IEEE Transactions on Visualization and Computer Graphics*, 27(7):3168–3181, 2021. doi: 10.1109/TVCG.2019.2963659 3