"How Do You Even Know That Stuff?": Barriers to Expertise Sharing Among Spreadsheet Users

QING (NANCY) XIA, University College London, UK

ADVAIT SARKAR, Microsoft Research Cambridge; University College London; University of Cambridge, UK

DUNCAN P BRUMBY, University College London, UK

ANNA L COX, University College London, UK

Spreadsheet collaboration provides valuable opportunities for learning and expertise sharing between colleagues. Sharing expertise is essential for the retention of important technical skillsets within organisations, but previous studies suggest that spreadsheet experts often fail to disseminate their knowledge to others. We suggest that social norms and beliefs surrounding the value of spreadsheet use significantly influence user engagement in sharing behaviours. To explore this, we conducted 31 semi-structured interviews with professional spreadsheet users from two separate samples. We found that spreadsheet providers face challenges in adapting highly personalised strategies to often subjective standards and in evaluating the appropriate social timing of sharing. In addition, conflicted self-evaluations of one's spreadsheet expertise, dismissive normative beliefs about the value of this knowledge, and concerns about the potential disruptions associated with collaboration can further deter sharing. We suggest that these observations reflect the challenges of long-term learning in feature-rich software designed primarily with initial learnability in mind. We therefore provide implications for design to navigate this tension. Overall, our findings demonstrate how the complex interaction between technology design and social dynamics can shape collaborative learning behaviours in the context of feature-rich software.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **Empirical studies in collaborative and social computing**.

Additional Key Words and Phrases: Spreadsheets; Knowledge sharing; Software learning; Workplace; Collaboration; Knowledge management; Thematic analysis

1 Introduction

For millions of users around the world, the spreadsheet is an essential tool for data handling and collaboration [70, 104, 106, 108]. Nearly 88% from a sample of around 1600 individuals reported working on spreadsheets shared with at least one other user [70]. Spreadsheet sharing activities, where viewing or editing rights are granted to another individual, are important for a variety of

Authors' Contact Information: Qing (Nancy) Xia, University College London, London, UK, nancy.xia.18@ucl.ac.uk; Advait Sarkar, Microsoft Research Cambridge; University College London; University of Cambridge, London, UK, advait@microsoft.com; Duncan P Brumby, University College London, London, UK, d.brumby@ucl.ac.uk; Anna L Cox, University College London, London, UK, anna.cox@ucl.ac.uk.

purposes. These include: ensuring data transparency, facilitating learning, enabling quality checks, and supporting teamwork [102, 106, 108].

However, spreadsheet collaborations can also be vulnerable to errors, the consequences of which are often costly [23, 30, 44]. In certain circumstances, the likelihood of such errors can increase. One common cause of errors is a mismatch in software-related expertise between collaborators [107, 108]. Two of the five most frequently reported spreadsheet errors occur when users fail to detect mistakes, or fundamentally misinterpret the final output of inherited spreadsheets [23]. Another is organisational knowledge loss, such as when experienced users move on from the organisation without having shared their knowledge to others beforehand [107]. This is particularly problematic in the spreadsheet context, since often only the original spreadsheet creator possesses the unique technical, domain, and contextual knowledge needed to maintain and interpret the spreadsheet [64, 108]. Addressing such knowledge loss is laborious, and the remaining users are often forced to develop workarounds rather than the expertise needed to fill in the gaps [107].

Developing users' technical skills is therefore crucial for reducing spreadsheet errors, particularly in professional contexts. While various learning approaches exist, we focus on knowledge/expertise sharing practices [28] among colleagues, such as over-the-shoulder learning [111]. Peer interactions are widely recognised as one of the most important ways to learn about software usage [61, 70, 85]. Furthermore, this practice ensures retention of essential, work-related knowledge within an organisation, even when employees transition or transfer [107].

However, designing interventions that effectively facilitate knowledge sharing is a complex problem. Sharing interactions are not only multi-directional, but are also inherently social, and can involve multiple parties [96, 118]. Typically, software-related expertise sharing is understood as a general form of knowledge sharing activity (e.g. answering questions) which happens to concern knowledge *about* a software (e.g. [58, 61]). Our understanding of such sharing activities are well supplemented by previous research in the field of knowledge management. These works provide key insights into how individual, social, and organisational factors (e.g. self-efficacy, social norms, rewards, and organisational culture) can affect expertise sharing behaviours [5, 99].

In this paper, we highlight that certain types of feature-rich software—such as the spreadsheet—are embedded with their own specific social norms, expectations, and associated identities, even if they have not inherently been designed to evoke these associations [101]. These software-specific associations can have an important influence on user's engagement with productive, collaborative behaviours. This is well known in cases where technologies have been explicitly designed for social interactions (e.g. social media platforms [21]), where unique sets of etiquette and taboos can develop and shape the way people engage in the platform. We suggest similar norms and expectations also exist for professional-use software such as spreadsheets, and can serve as social barriers to expertise sharing practices. For example, expectations of ownership can affect *when* and *to what extent* it is acceptable to make changes in another user's spreadsheet [27]. Similarly, beliefs about what distinguishes a spreadsheet 'expert' or a professional programmer from an 'amateur' can affect views on *who* is allowed to critique another individual's spreadsheet [103].

We argue that paying greater attention to the socially constructed perceptions surrounding software expertise and acceptable usage practices can deepen our understanding of why people may, or may not, engage in expertise sharing. We apply this perspective to the spreadsheet context, with the goal of generating insights into the following questions:

RQ1: What are the social barriers to expertise sharing among spreadsheet users?

RQ2: How does the design and usage of spreadsheet systems contribute to the social dynamics of expertise sharing between users?

In our paper, we focus on spreadsheet sharing behaviours—defined as granting another user with either physical or digital access to one's spreadsheets for viewing or editing purposes. This behaviour was selected because research suggests that spreadsheet sharing can facilitate expertise sharing both explicitly and incidentally. In other words, while the primary goal of sharing may be collaboration, opening access to a spreadsheet often allows recipients to uncover new features and embedded practices, and therefore has important implications for learning [102].

Our work provides two main contributions to the CSCW community. First, we offer a detailed description of the key factors which deter spreadsheet expertise sharing. Through our interviews, we found that spreadsheet providers often struggle to adapt highly personalised use strategies to subjective standards of quality and to navigate the social timings for sharing. These challenges are compounded by concerns over potential collaboration errors, conflicted self-evaluations of expertise, and broader normative beliefs that devalue spreadsheet-related knowledge. Altogether, these findings highlight how feature-rich software amplifies the social challenges of sharing due to their inherent flexibility, technical complexity, and potential ambiguity in how contents and functions are perceived.

The second contribution of our work is a set of design suggestions which aim to mediate the aforementioned challenges and promote knowledge sharing in the spreadsheet context. We link our observations to previous discussions around designing feature-rich software for ease of learnability, and build on previous literature to highlight how alternative design choices may lead to different user dynamics and engagement in collaborative behaviours in the long term.

2 Related works

2.1 Spreadsheets as 'praxisware'

Designing for software learnability has long been a core topic of interest in the field of Human-Computer Interaction [42]. While the field typically focuses on initial adoption and accessibility, designing for long-term learning in feature-rich software presents unique challenges [42, 101]. We borrow the term 'praxisware' from Sarkar [101] to discuss the types of software for which the issue of long-term learnability is particularly salient. Praxisware are distinguished from general-use software in three ways: 1) the richness and diversity of functionalities which they provide; 2) the longevity of their usage—typically spanning years or even the entire length of an individual's career; and 3) being deeply embedded within specific professional identities and communities, unlike general-use software [101]. In other words, they are tools which both shape and allow users to express their professional skills, values, and identity (e.g. Photoshop, AutoCAD) in a way that is not encapsulated in other types of software (e.g. Microsoft Word, PowerPoint)[90]. This incentivises learning, as development of one's technical expertise is not only practical, but can also have important personal, social, and professional implications [71].

In this paper, we consider spreadsheets to be a type of praxisware. Firstly, spreadsheets are feature-rich, offering a wide range of functionalities that can be challenging to master without long-term investment and support from others [48, 102, 107]. Secondly, spreadsheets are strongly affiliated with professions such as finance, economics, and banking [30, 70], so high levels of spreadsheet expertise are often closely associated with high levels of professional expertise [90]. This means the study of spreadsheet-related knowledge sharing will have implications for both software learning and professional knowledge sharing contexts. Finally, previous works suggest that a 'spreadsheet user' is in itself a marked social identity [18] with specific social associations and traits distinct from related identities such as 'programmers' [48, 103]. This suggests that social norms and professional implications related to spreadsheet learning and expertise could influence engagement with knowledge sharing practices. By studying spreadsheet software, we can provide

valuable insights into how attitudes, norms, and social identities around spreadsheet use could influence collaborative behaviours.

2.2 Supporting software learning through expertise sharing

Colleagues play a crucial role in supporting expertise development in software use [61, 71, 111]. Studies show that learning from colleagues through conversations and demonstrations happen more often and is more preferred over alternative resources such as search engines, training courses, help pages, or videos [61, 70, 71].

Given the prevalence of interpersonal learning, researchers have proposed various technical innovations to help users better leverage colleagues' expertise for learning purposes. A key tension which researchers must consider is the balance between the depth of expertise shared, and the workload of the helper [40, 69]. Some tools prioritise giving user access to experienced individuals when needed. For example, expert recommendation systems facilitate searches for knowledgeable individuals in an organisation by generating profiles about other users [1, 55, 92, 98, 105]. Tools such as 'MarmalAid' [26] and 'MicroMentor' [58] support the communication process by providing easy access to individual experts for synchronous in-context help. Other approaches focus on reducing the effort associated with capturing and documenting knowledge for sharing (e.g. [25, 43, 65]), such as by using crowd-sourcing approaches (e.g. [20, 34, 65, 66, 79, 80]). The goal is to make knowledge sharing interactions more easily accessible and less effortful for sharers, therefore increasing possible learning opportunities for recipients.

While the above research helps to identify possibilities for how expertise sharing activities can be supported, previous studies often implicitly assume that knowledgeable individuals are always willing to share or reveal their expertise. However, there are also many undesirable social repercussions associated with publicising one's expertise in the workplace [40, 69, 96]. For example, individuals who are deemed 'more' experienced than others may be disproportionately allocated responsibilities or expected to meet unrealistic demands [32, 116]. Giving other users visibility over one's software usage metrics or a means to quantify their own expertise can lead to concerns of surveillance, fears of criticism, and social comparison with others [40, 96]. Additionally, as Pipek et al. [97] point out, in professional contexts, the role of the learner and helper are not mutually exclusive or binary, and one person often occupies the role of both interchangeably depending on the context. However, if an individual is unable to transition from a learner to helper, or recognise how their expertise may support others, then they are unlikely to engage in knowledge sharing activities—even if they are sufficiently equipped.

The effort of sharing is therefore not the only barrier to knowledge sharing. Identifying the social and individual consequences of sharing, from the knowledge sharer's perspective, is also essential. Our goal is to inform the design of spreadsheet technologies based on that perspective.

2.3 Knowledge management and contributors to knowledge sharing

'Knowledge management' is the study of how knowledge, as a type of intellectual asset, should be retained and cultivated within organisations [10, 57]. An integral process in the management of these intellectual assets is the process of knowledge/expertise sharing [10, 113]. There are two common ways in which the process of sharing is defined [1, 78]: 1) sharing 'knowledge' through artefacts, by explicitly formalising knowledge through documentation or storing in an accessible database; 2) sharing 'expertise' via interpersonal communication, which supports the transfer of information that is inherently challenging to document (e.g. personal and practical experience, intuition, know-how [89]). As both documented knowledge and tacit expertise are relevant for feature-rich software operation, we use both terms interchangeably in this paper.

Existing research from knowledge management has sought to provide a comprehensive understanding of the motivating factors influencing knowledge sharing intention and behaviour [36, 113]. Theoretical frameworks of behaviour and decision-making, such as the theory of social exchange and social capital [120], social cognitive theory [8], or the theory of planned behaviour (TPB) [2, 3], are frequently applied in survey-based studies to support researchers in identifying the relationships between relevant determinants of sharing behaviours (for reviews, see [36, 110, 113]).

Such studies establish an important foundation for understanding the processes underlying knowledge sharing in the workplace. For example, cultivating cultures which establish pro-sharing norms [15, 59], fostering a collaborative rather than a competitive environment [115], and reducing the emphasis on rank and hierarchy [84] can facilitate information flow between individuals and groups. In contrast, where knowledge sharing may be viewed as a loss of an individual's own unique expertise or asset [109], or where there may be challenges associated with impression management (e.g., if knowledge providers are seen to be self-serving or disagreeable through the act of sharing) [16], knowledge sharing intention is often lower. Interpersonal relationships, such as the degree of trust between individuals [59], as well as an individual's own characteristics, such as self-efficacy [14, 53], can also affect a person's willingness to engage in sharing activities.

Another key focus in knowledge management is the use of technologies to facilitate the *communication* of knowledge. Previous works have applied frameworks such as TPB to explore how online/virtual platforms and social media facilitate knowledge sharing activities (for reviews, see [35, 36]). These studies demonstrate that the quality of a system (e.g., its ease of use, its accessibility) can positively affect sharing behaviours by increasing users' perceived behavioural control (i.e. their belief in their ability to execute a desired behaviour [2, 3]) and their perception of resource availability (e.g. feeling like they have more time available) [52, 73]. However, while effective system designs can be useful in producing positive sharing experiences, embedding the social nuances and motivators effectively into technological systems can be challenging. For example, implementing virtual social rewards (e.g. social points for answering questions on a forum) can, in some cases, undermine the positive feelings of altruism, enjoyment in helping others, and desire to build connections, which otherwise predict knowledge sharing on these platforms [112, 121].

It is therefore essential that technology designers and researchers clearly understand the social influences underlying interactions among users if we are to design systems that effectively promote expertise sharing. To contribute to that understanding, our study aims to identify how associated norms in spreadsheet-specific practices influence expertise-sharing intentions.

2.4 Opportunities and challenges for expertise sharing in spreadsheet collaborations

Several prior works have explored knowledge sharing activities around spreadsheets. Research suggests that, alongside numerical data, a spreadsheet records the traces of how an analysis is conducted, the history of decision-making and intentions [24, 64, 108], as well as contextual labelling and interpretation of information [60] (e.g. 'Department Budget 2024'). Furthermore, spreadsheet sharing triggers interpersonal communication, encouraging users to externalise tacit knowledge via annotations, comments, and edits within or around a spreadsheet file (e.g. via in-person help-seeking, emails, or meetings) [24, 64, 102, 108].

Research focusing on spreadsheet has explored the challenges and opportunities for spreadsheet-related knowledge sharing. One key finding is that internal organisational learning from peers is particularly valuable as users are not always at liberty to share confidential information in spreadsheets outside the organisation [107]. Conversations with other users and help-seeking interactions are also commonly reported methods for self-learning among spreadsheet users [70, 102]. Indeed, the absence of these interactions and a failure to transfer knowledge to relevant parties can result in severe knowledge gaps within the organisation which can negatively impact

performance [107]. It is therefore important to develop a greater understanding of the social processes which may deter users from effectively utilising these opportunities for learning.

Prior research suggests that the technical complexity required to find and interpret information in spreadsheets can be a barrier to such forms of knowledge transfer, unless users directly communicate and clarify with each other [108]. Common technical issues exacerbating the problem include: the issue of visibility and presentation—users must interact directly with cells in order to reveal the underlying formulae or to structure the data in a way which is meaningful [9, 24, 108]; issues of logic interpretation and implementation—particularly when confronted with logically complex or visually dense formulae [50, 51, 91, 107, 119]; or issues with the language of the interface—which are often based on spreadsheet-terminology (e.g. 'cell A1') rather than representations clear to users or their specific domains (e.g. 'profit,' 'budget') [31, 47, 50, 60].

Findings of spreadsheet research suggest that documenting and extracting tacit knowledge may be too effortful. Users may therefore perceive the required resources (e.g. time, attention) as too high for knowledge sharing to be worthwhile. However, to our knowledge, there has been little to no explicit focus in the current research on the social norms, attitudes, and beliefs around spreadsheet-related knowledge sharing and why users may or may not choose to engage in these practices. Our study aims to address this gap.

3 Method

To address the research questions previously stated, we analysed the interview data collected in two separate studies, originally designed for another research exploration. We decided to use this data to investigate the current questions for two reasons. First, interviews in both studies included sections dedicated to topics such as spreadsheet collaboration, help-seeking and help-giving behaviours, practices used to support spreadsheet comprehension when sharing, and discussions of industry standards and social norms of spreadsheet use. Second, interview sets from each study offered complementary perspectives advantageous for triangulating the findings. Study 1 provided a general overview of the social, organisational, and industry-based contexts surrounding spreadsheet use and learning. Study 2 provided insights into how users applied, discovered, explored, and communicated about different aspects of the spreadsheet (with a specific focus on formulas) when working with others.

In the following sections, we broadly outline the original aims, recruitment strategies, and materials of each study. We place our emphasis on materials directly relevant to the questions in this paper. Ethical procedures of these studies are included in Appendix A.

3.1 Study 1

3.1.1 Original aims. Study 1 was conducted between February and March 2023. The original goal was to explore how users' interactions, practices, learning, and mental models of the spreadsheet differed according to their professions. It was therefore important that participants were recruited from a variety of professional backgrounds.

To ensure sample diversity, the chosen professional domains from which we invited participants had to be sufficiently distinct in terms of their spreadsheet practices and resources for spreadsheet learning. For that, we evaluated the domains based on five key criteria that we identified:

- *Organisational structure* (e.g. whether the nature of an individual's role is team-based or predominantly driven by the individual);
- *Prevalence of spreadsheet use* (i.e. the estimated proportion of time dedicated to spreadsheet-related work within working hours);

Table 1. Demographics of Study 1 interview participants

ID	Age (Gender)	Occupation	Seniority	Self-reported spreadsheet expertise
Technical domain				
S1P5	18–29 (M)	IT consultant	Non-manager	Extensive experience, some expertise
S1P32	40–49 (M)	Data migration specialist	Supervisor/ Manager	Some experience, still a beginner
S1P39	40–49 (M)	Self-employed IT tester	Non-manager	Extensive experience, some expertise
S1P15	60+ (M)	Independent app developer	Executive	Extensive experience, some expertise
Financial/Data science domain				
S1P1	18–29 (M)	Health economist	Non-manager	Very experienced, high expertise
S1P3	18–29 (M)	Consultant	Non-manager	Extensive experience, some expertise
S1P4	18–29 (F)	Data scientist/ Researcher	Non-manager	Extensive experience, some expertise
S1P21	40–49 (F)	Head of Treasury	Supervisor/ Manager	Very experienced, high expertise
Science domain				
S1P26	18–29 (M)	Biomedical researcher	Non-manager	Extensive experience, some expertise
S1P29	40–49 (F)	Biochemistry researcher	Supervisor/ Manager	Extensive experience, some expertise
S1P34	18–29 (M)	Neuroscientist	Non-manager	Some experience, still a beginner
S1P45	18–29 (M)	Chemistry research assistant	Non-manager	Some experience, still a beginner
Administrative domain				
S1P2	18–29 (F)	Sales and distribution	Non-manager	Some experience, still a beginner
S1P22	40–49 (F)	Auditor	Supervisor/ Manager	Some experience, still a beginner
S1P23	18–29 (M)	Marketing consultant	Executive	Extensive experience, some expertise
S1P24	30–39 (F)	Human resources	Supervisor/ Manager	Extensive experience, some expertise
S1P25	30–39 (F)	Facilities manager	Supervisor/ Manager	Extensive experience, some expertise

- *Required spreadsheet expertise for domain* (the expected difficulty, commitment, and level of knowledge and mastery of spreadsheets necessary to carry out daily tasks in the field);
- *Relevance of field-specific knowledge to spreadsheet use* (e.g. an individual possessing knowledge of programming languages may be more proficient at using spreadsheet formulas compared to an administrator);

- *Prevalence of organisational support for spreadsheet use* (e.g. the availability of training sessions, workshops, guidelines, online or academic courses, informal support or 'buddy' systems etc

We ultimately identified four professional domains (technical, financial/data science, scientific, and administrative) to recruit participants. A detailed evaluation of these domains in terms of the five criteria is included in the supplementary materials.

3.1.2 Participants. We identified participants in these four selected domains through personal connection, professional social networks (e.g. LinkedIn), and a pre-study survey distributed on the online participant recruitment platform Prolific. Using Prolific allowed us to extend our recruitment beyond the researchers' immediate circles. Though the use of online platforms can pose risks to the validity of participant responses, previous reviews suggest that participants on Prolific platforms overall demonstrate high engagement and attention to detail [33, 93].

All participants filled out a 10-minute pre-study recruitment survey, which allowed us to verify and ensure diversity in their professional background, spreadsheet usage, and demographics, before they were invited to the interview. Overall, 62 participants filled out the survey, 17 of whom then proceeded to the interview. Four participants were recruited for three of the four professional domains. For the administrative domain, five participants were recruited. The demographics of these 17 participants are shown in Table 1.

3.1.3 Materials. The study involved a pre-study survey for recruitment and an interview guide. The survey was a shortened form of the Spreadsheet Engineering Research Project (SERP) survey, a well-established tool used to summarise spreadsheet practices in organisations [6, 70]. We condensed the SERP survey from its original 67 questions to 20, as our main goal was to capture broad overviews of participants' spreadsheet practice which we then used to generate in-depth follow-up questions during the interviews. We therefore did not want participants to dedicate excessive amounts of time to the survey. We selected one or two relevant questions from each section of the SERP survey, focusing on participants' demographics, professional backgrounds, self-reported spreadsheet expertise, purpose of spreadsheet use, time spent on spreadsheet-related work, size and complexity of spreadsheets, number of spreadsheet collaborators and sharing practices, organisational policies, and frequently used spreadsheet learning resources. Topics which were deemed less relevant for recruitment purposes and more suited to interview explorations (e.g. spreadsheet risk management) were excluded. We also added 6 questions exploring whether participants are self-taught as well as their familiarity with a range of spreadsheet features, experience with programming languages, and willingness to participate in the interviews.

The interview guide comprised two sections. The first section was a general discussion addressing three main areas. The first area was about clarifying survey responses and gathering information on participants' work responsibilities and spreadsheet usage. The second examined spreadsheet-related practices, including team collaboration and spreadsheet-sharing habits. The final area focused on spreadsheet learning, user expertise, and help-seeking practices.

The second section of the interview guide prompted users to share and briefly describe an example spreadsheet they used at work. We first asked questions about the specific application of features for this particular spreadsheet (e.g. the use of formulas, data layout, visualisation strategies). We then asked participants to explain the decisions behind the creation of the spreadsheet. Participants were then encouraged to reflect on the challenges they faced in spreadsheet use and to provide suggestions for future tools to address these challenges (see the supplementary materials for the Study 1 survey and interview guide).

3.1.4 Procedure. A survey link was shared with participants on our social networks and Prolific. Participants who filled out the survey were invited on a rolling basis, and they booked their own

Table 2. Demographics of Study 2 participants

ID	Domain	Self-reported spreadsheet expertise
S2P1	Spectrum systems engineer	A lot of experience, and I use some advanced features
S2P2	Data scientist	A lot of experience, and I use some advanced features
S2P3	Business operations specialist	A lot of experience, and I use some advanced features
S2P4	Teaching administrator	A lot of experience, and I use some advanced features
S2P5	Commercial real-estate finance analyst	A lot of experience, and I use some advanced features
S2P6	Machine learning & data science student	A lot of experience, and I use some advanced features
S2P7	Financial analyst	A lot of experience with many advanced features
S2P8	HCI PhD student	A lot of experience, but my use is basic
S2P9	Lawyer/MBA student	Some experience, but I'm still a beginner
S2P10	Data scientist	A lot of experience, but my use is basic
S2P11	Product manager	A lot of experience, and I use some advanced features
S2P12	Centre co-ordinator	A lot of experience, but my use is basic
S2P13	Civil servant	A lot of experience, but my use is basic
S2P14	Statistics student/ex-MBA analyst	A lot of experience, and I use some advanced features

timeslot. The interviews lasted around 1 hour, and were all conducted, recorded, and transcribed online by the first author using Microsoft Teams, an online video meeting platform, and its built-in transcription services. The order of questions were kept flexible to ensure a natural flow around the participants' responses. Participants were first asked a set of follow-up questions based on the survey. They were then brought into the interview sections for an in-depth discussion. Finally, they were invited to screen-share and discuss their example spreadsheets.

3.2 Study 2

3.2.1 Original aims. Study 2 was originally conducted between August and October 2018. It focused on understanding how users interact with spreadsheet formulas, which are an important aspect of spreadsheet use, and how they developed expertise in using them. Because the study prioritised interaction, learning experience, and the factors affecting users' engagement with spreadsheet formulas, recruitment was based on diversity in self-reported spreadsheet expertise and programming expertise.

3.2.2 Participants. A total of 14 participants were recruited from the research organisations and partners we are affiliated with as well as from a pool of participants engaged in our previous research projects. The demographics of these participants are outlined in Table 2.

3.2.3 Materials. Similar to Study 1, Study 2 also involved a short pre-study survey for recruitment, and an interview guide. The survey captured the key demographics of the participants, including occupation, self-reported spreadsheet experience, spreadsheet use purpose, as well as expertise in both spreadsheet formulas and programming languages.

The design of the interview guide focused specifically on the barriers encountered in using spreadsheet formulas and the learning process towards achieving mastery in it. It consisted of four broad areas: discussion of participants' background (i.e. their work, daily tasks, and work tools);

how they developed their expertise in spreadsheet use, and the facilitators and barriers to their general spreadsheet use experience.

In addition, the interview guide consisted of a spreadsheet walk-through segment where participants provided example formulas, followed by an interactive segment, where they demonstrated how they would author one of those and compare the same workflow in Google Sheets and Microsoft Excel. Of particular relevance to our current research question was that the guide was designed to identify the specific spreadsheet practices and features used during collaborative work. In that sense, the Study 2 interview guide was supplementary to the Study 1 guide, which focused on abstract social contexts and influences underlying collaboration (See the supplementary materials for Study 2 interview guide).

3.2.4   Procedure. Participants were invited via email and asked to fill out the survey capturing their demographics. Interviews were between 1-1.5 hours long, conducted in-person, and recorded via audio recorders or Microsoft Teams. Each interview began with a discussion around the participant's background and how they developed their knowledge of spreadsheet formulas. Following the discussion, each participant was asked to walk us through their example formulas and demonstrate how they write or apply those to their work.

3.3   Analysis

3.3.1   Positionality statement. In data analysis, we adopted critical realism as an epistemological position [17, 95], and accordingly, we interpret the participants' language as a reflection of their unique, personal experience of spreadsheet use, collaboration, and learning [46]. This is important as we were interested in the underlying beliefs, attitudes, and social norms that obstruct the willingness to engage in expertise sharing—but recognise that the participants themselves may not describe these with the same terms we do.

3.3.2   Theory of planned behaviour (TPB). We used the theory of planned behaviour (TPB) to interpret the findings of Study 1 and Study 2. TPB explains social-related behaviours as intentional acts of decision-making, informed by three factors: attitude (expected outcomes of the behaviour), social norms (external expectations and motivations to conform to these norms), and perceived behavioural control (consisting of internal self-efficacy, and external constraints such as resource availability) [2, 3]. TPB is therefore well-suited for providing a comprehensive overview of volitional behaviours, such as knowledge sharing, and has been widely applied in knowledge management research for this purpose [36, 87]. For both of our studies, TPB helped to identify the individual, social, and technological factors influencing the intention to share spreadsheet-related expertise.

3.3.3   Thematic analysis. Both Study 1 and Study 2 interviews were treated as a single, complementary dataset, and analysed using reflexive thematic analysis. We applied both inductive and deductive approaches to qualitative coding by utilising the 'template analysis' method established by King [63]. 'Templates' (i.e. pre-defined code structures) related to the core research goals were generated and applied a priori with the flexibility to expand with new codes as the analysis progressed. This was particularly useful as we needed to organise the coding template under TPB, while also ensuring that the additional codes were reflective of experiences and beliefs specific to the spreadsheet context.

We began by inductively coding the participants' descriptions of their interactions, beliefs, and experiences of using and sharing spreadsheets. The initial work of developing and refining the coding template was carried out by the first author. To ensure analytical rigour and reflexivity, the final template and a selected sample of quotes were shared with the second author, a senior researcher

with prior experience in spreadsheet research. Their domain expertise provided alternative insights and served to triangulate the first author's interpretation [45].

Once an agreement was reached between the authors, we followed a deductive approach and reorganised the codes according to the components of the TPB framework: attitudes around spreadsheet-related expertise sharing, subjective norms of spreadsheet-related expertise sharing, and perceived behavioural control over spreadsheet-related expertise sharing. The supplementary materials includes documentation of how the coding template was developed and refined over the course of analysis, and how agreement was reached between the two authors.

4 Results

In the following sections, we present our findings under themes by incorporating quotes from the Study 1 and Study 2 participants denoted S1P(no) and S2P(no) respectively. We found that the participants have competing attitudes towards sharing activities, and they deliberately determine when and how to share spreadsheet-related knowledge. Furthermore, these decisions are often influenced by perceived effort of communication, considerations of social timing and roles, conformity to perceived norms, and evaluation of one's own expertise and resources.

4.1 Attitudes around spreadsheet-related expertise sharing

4.1.1 Spreadsheet recipients as a source of errors and disruption. Participants had different, often conflicting attitudes towards spreadsheet sharing behaviour, depending on their role as providers or recipients. Among providers, spreadsheet sharing was commonly associated with frustrations, confusion from recipients, and an increased likelihood of errors. Recipients were often described as potential sources of changes disruptive to providers' goals. For instance, S2P3 noted: "*Because all the stuff I do is touched by so many people [...] I don't want them to mess with it.*" Participants who created their own spreadsheets anticipated that other users may struggle to understand and correctly implement changes according to their original intentions. An example is S1P4, who stated that "*you've set up a bunch of different formulas and you know how they work and why they work. Someone else might not know.*"

As a result, participants intentionally limit input from others to minimise possible disruptions. Some restrict user rights, lock files, or even socially enforce their role as the sole editor or controller of the spreadsheet from the start. Continuing with their example, S1P4 provided an example of how they would protect a sheet that was crucial to their calculations by "*lock[ing] the entire Excel file, because it's going to be used by a couple of people [...] if they accidentally hit something then things could break and it takes a long time to figure out why it's not working.*"

4.1.2 Knowledge sharing interactions provide users with significant support. Despite the challenges faced by providers, recipients consider these individuals useful sources of knowledge. Solutions, methods, and examples found in others' spreadsheets or in answers online could easily be reused for recipients' own purposes. S1P1 described such an instance where they took features from their colleagues' model: "*So largely, I have used examples from other models. I've kind of patched some things together because some things work better than others.*" In some cases, recipients may inherit spreadsheet templates as examples from other colleagues, which they then develop according to their own preferences (S1P2, S1P26).

The participants further acknowledged the usefulness of knowledge sharing interactions when they are the recipients. Providers' advice and suggestions often prompted further discovery and, as was the case for S2P13, even increased motivation for learning: "*There was a particular colleague who [...] just said something offhand about how much easier this would be to do in Excel, and from there, that kind of prompted me to then start asking questions of other people who I know use it, including*

this guy, to start using it more." S1P3 also reported that colleagues may re-share discoveries of useful features with others in their local community (e.g. via an internal Microsoft Teams channel).

Some participants also engaged in proactive help-seeking from others, though this did not always lead to learning. 'Local' providers (i.e. those with personal associations to the individual) were typically viewed as trusted sources of expertise, as S2P14 exemplified: "*Sometimes there will be more senior bankers sat around you that you go up to [to] ask, 'I've got this problem. I've never encountered it before. In your years of experience, have you come across anything like this?'*" However, some participants developed more dependent relationships. In these instances, rather than developing their own skills, they may instead delegate more challenging tasks to others for support. For instance, S1P32 described the following: "*I basically go, 'Ooh, can't do this,' and I've fired it down for the spreadsheet person. They'll fix it for me, you know?*" S1P25 similarly explained that, when they encountered information which they knew was "*above [their] head, [their] skill set, and [their] understanding,*" they would "*[do] all the bits [they] needed to do […] and then ask somebody else to actually help [them] pull it through.*"

4.2 Subjective norms of spreadsheet-related expertise sharing

4.2.1 Familiarity with spreadsheets thought to be commonplace, downplaying the importance of expertise. Participants frequently expressed a belief that 'most' individuals would be familiar with using the spreadsheet, prompting unfounded assumptions about others' spreadsheet expertise. These views are largely driven by the intuitive nature of spreadsheet interactions, and a perception that spreadsheets are widely available and well established. For instance, S1P32 stated: "*Look, Excel, everybody knows how to use it […]. You just have to right click and add a new tab, and you're pretty much done.*" These beliefs often led participants to make assumptions about others' expertise, which often led to surprise when these assumptions are violated: "*I'm always surprised when people don't know how to use formulas in Excel or automatically expand things or do conditional formatting*" (S1P26). Some participants, such as S1P24, expected greater proficiency from younger colleagues, and felt surprised when this was not the case: "*[It's] really surprising to me. A lot of younger people leaving school don't seem to be able to use Excel.*"

The belief that 'most' people possess some form of spreadsheet expertise has negative implications when reflected in organisational structures. Notably, it appears to close down opportunities for spreadsheet-related knowledge sharing. For example, S1P22 noted that they were unable to get answers to their queries about Microsoft Excel from their Information Technology (IT) department, jokingly stating that the department found their request to be out of their responsibility and "*beneath them.*" An extreme example was given by S1P24, who, despite lacking technical knowledge, was forced to recover corrupted spreadsheet data from a virus because their organisation wanted to avoid outsourcing the issue to an external IT team.

Dismissive attitudes towards spreadsheet education, both in industry and in academic institutions, directly contribute to challenges in recruitment. S1P32 reflected: "*When you find that person who knows Excel, they're your best friend for a lot of the day. But it's never something we look for when we're interviewing, bizarrely.*" S1P21, who works in a treasury, similarly noted: "*They [students] don't actually use Excel in [university] ever. […] it would be impossible to recruit students from university level who have the skills you need, it'd be 1 candidate out of 10,000.*" As a result, spreadsheet training for new recruits (e.g. S1P21) is either the responsibility of existing managers or employees, or else recruits are simply asked to learn on the job. This, as S1P3 (a junior consultant) reflected, becomes "*a little overwhelming at first, especially when you're thrust in a work environment and they throw a bunch of numbers at you and they're like 'hey, do your best!'*"

In some cases, even expressing an interest in learning and using spreadsheets seems to trigger dismissive sentiments and negative judgments. The participants' statements indicate that it was

socially acceptable to question those who dedicate time and effort to learning about and using spreadsheets. S1P29, a chemist, remarked: "*I think we probably think of them [spreadsheets] as being a necessary evil. I do think there's some people that love them a bit too much.*" Further, even when positively recognising a colleague for solving a problem, S2P9 used the label "*Excel nerd.*" Individuals who expressed interest or excitement at the introduction of new spreadsheet functionalities could encounter surprise or skepticism from others. This was the case for S1P21: "*When they introduced XLOOKUP [...] I was so excited and all my colleagues were like: 'why? what's wrong with you?'*"

4.2.2 Personalised spreadsheet practices are valued, but there are pressures to conform to poorly defined 'normal' practices. Participants frequently described the spreadsheet's capacity to support personalised sensemaking practices as a great asset. Relatedly, participants acknowledged that the way spreadsheets are used varies across individuals. "*Every senior banker has a different style, as well, and they'll ask for deviation from the template,*" S2P14 commented when describing their work with senior colleagues. Participants also discussed their own personalisation needs and appreciation for freedom and control that the spreadsheet offers. S1P22 likened the spreadsheet to "*a big empty notebook [they] can use to [their] will.*" Similarly, S2P4 stated that, when using a spreadsheet, they can "*organise everything in a way that makes the most sense to [them].*"

Despite these advantages, participants' accounts suggest that spreadsheet sharing in practice is often subject to certain expectations of standards and presentation—though the expectations may not be explicit. "*There's guidelines around what the data needs to look like, but not guidelines necessarily of how you need to do that and make that work,*" said S1P25 when describing the types of data that their organisation expected to receive from clients. An awareness of these implicit expectations, and a desire to conform—or appear to conform—to these standards, mean that the personalised elements of participants' spreadsheets often pose a barrier to sharing. For example, S2P10 stated: "*I wouldn't even really want to share the template for it, if the company asked me to provide it, because I feel this is very... it suits my needs.*"

4.2.3 Sharing may not be appropriate depending on timing and individual roles. Even in a collaborative context, individual roles and responsibilities are typically varied. Some users have rigid expectations as to who needs to complete a particular task. As a result, collaboration is typically sequential, in that users carry out work only *after* the previous user has completed theirs, rather than simultaneous, real-time discussion and collaboration. For instance, S1P1 described: "*Generally we avoid making changes [on others' spreadsheets] because if you have too many people doing things then you can lose track of how things are done.*" This was also illustrated by S1P5, an IT consultant, describing the process of data cleaning with their team: "*They're [the colleagues] waiting for whatever data to come through, so I do need to be aware of that.*"

The extent to which individuals feel they can rely on others is also influenced by their own roles and identity. For example, S1P25 felt compelled by their seniority to develop proficiency and independence: "*As a head of department [...] there's things I can go and say to someone like the Data and Insights team, 'I need this,' [...] but if I'm expecting my account managers and people working underneath me to be able to do these things that I feel I have to demonstrate it myself as well.*"

Some participants were restricted by data confidentiality and privacy policies in terms of which spreadsheets they could share or what aspects of their work they could discuss with colleagues. "*Historically, [...] I would have been very cautious about who got a copy of a particular spreadsheet,*" S1P15 noted. As a consequence, exposure to others' spreadsheet practices or feature usage is limited, as S1P4 explained: "*I don't know if I know of how other people in my industry still store spreadsheets, [...] we don't share any of this stuff.*"

4.3 Perceived behavioural control over spreadsheet-related expertise sharing behaviours

4.3.1 Users tended to negatively evaluate their spreadsheet expertise. Confidence in one's own knowledge is an important prerequisite to engage in knowledge sharing. However, some participants' statements showed they lack such confidence. For instance, S2P12 stated: "*I can do basic formulas, but I'm not, you know, doing pivot tables and all that kind of thing.*" Some participants were overwhelmed by the extent of the spreadsheet software's features, assuming that they only accessed a fraction of its capabilities. S2P13 speculated: "*There could be something out there that helps me with something, I'm just not aware of it.*" Similarly, S2P3, when asked to report on their own expertise, stated: "*Just the fact that there are books of them means there's a million [formulas] out there that I know nothing about.*" Since the participants commonly believe that spreadsheet expertise should be widespread (see section 4.2.1), a lack of expertise is portrayed in a negative light. An example is S1P32, who described their own knowledge as "*embarrassing,*" stating that "*you'd think after 20 something years my experience should be a little bit higher than it is.*"

Even the participants with confidence in their own knowledge—typically from finance or data science—noted with caution that their knowledge is restricted to their specific role. S2P10, a data scientist, stated: "*At the level I think I need to be at, I'm very confident with what I do.*" S1P2, who works as an economics consultant, stated: "*I feel I have very basic knowledge and specific knowledge of what I need to know, but aside from that, I would be clueless to do some basic things.*" This attitude stems from a lack of standardised definition for spreadsheet expertise, leading the participants to gauge expertise via various measurable attributes. These attributes range from comparisons with others' recall, speed, complexity of spreadsheet formulas to attributions based on individual characteristics such as age and profession. S2P3, for example, was impressed at their colleague's ability to memorise formulas: "*[They] literally [go] to a cell, and [they type] the whole formula [...] And I'm like, 'how do you even know that stuff?'*" Similarly, S2P5 discussed how "*younger guys in [their] office do a lot more funky stuff than [they] do.*" In some cases, individuals may classify the same spreadsheet features under different levels of difficulty, as S1P21 described: "*To them [new recruits], 'advanced' means pivot tables. And I would classify pivot tables and SUMIF, things like that, as basic.*"

This lack of clear standardisation led some participants to downplay the value of some beneficial spreadsheet practices. For example, when asked about their motivation for learning Visual Basic for Applications (VBA) and automating the functions within their spreadsheet, S1P21 stated that it is "*mainly because by nature [they are] actually quite lazy.*" As a result, S1P21 is motivated to find "*any quicker, easier way of doing [things].*" S2P13 similarly described their strategy of reusing spreadsheet templates or existing formula solutions online as "*laziness.*" Despite these points, both participants had self-reported high levels of experience in our surveys, suggesting that judgments of the value of spreadsheet expertise do not always correspond to the perceived depth of expertise.

4.3.2 Effort is required to support effective communication with others. When sharing, reviewing, and editing spreadsheets are unavoidable, participants often expend additional effort to modify, extract, and translate information in ways suited to the needs of others. In cases where a spreadsheet is complex, not all of the information in it is relevant for others. For instance, as S1P3 exemplified, "*a lot of times, they'll [clients] be speaking to senior leadership and they don't have the time to look at your model, like never.*" As a result, some participants were averse to the idea of excessive sharing, preferring to provide summaries of the problems they encountered or the spreadsheet's overall output instead. For example, S2P13 shared an example of a department budget spreadsheet, which they "*show[ed] to [their] director, so [they] don't include a lot of the other, smaller transactions.*"

Other participants noted how they would intentionally adapt their personal practices when anticipating that their spreadsheets would be shared with others. One experienced user, S1P21, explained how they would refrain from their usual practices during collaborations to accommodate others' expertise: "*If I have a spreadsheet that I know is just for myself, I do tend to do it properly [...] whereas if I do the full belts and braces, [other users] get really stuck because there's too much for them to unpick.*" S1P45 similarly described how they would be more concerned with formatting, centering, and aligning their inputs when sharing with others compared to using spreadsheets on their own.

In an extreme case, spreadsheet sharing was avoided altogether because the participant did not want to spend time explaining or cleaning data for others. S2P4, a teaching administrator, explained: "*I don't really share [spreadsheets], it's not that I don't want to, it's [...] because then they [colleagues] would say: 'why did you do this?' And I just can't spend time explaining to them the reasons of why I do everything, because it would just be too time-consuming.*"

5 Discussion

Overall, our findings reveal unique social barriers that spreadsheet users face in expertise sharing which require them to spend significant effort to address. While some of these barriers are driven by concerns of errors or miscommunication, we found that subjective standards of spreadsheet expertise and the timing of when sharing is considered appropriate also play a part. The design of spreadsheet software contributes to the participants' mixed self-evaluations of expertise and reinforces the assumption that spreadsheet skills are widespread, easy to acquire, and not explicitly valued in organisations. All these insights provide a comprehensive perspective of how sociotechnical factors influence spreadsheet expertise sharing.

In the following sections, we discuss how design aspects of spreadsheet software contribute to social norms influencing users' knowledge sharing intentions. We contextualise our findings within previous literature on spreadsheet-based collaborations, and propose implications for design to promote greater knowledge sharing engagement.

5.1 Spreadsheets as a platform for collaboration and knowledge sharing

Our research challenges some of the pre-existing notions of how work is distributed between spreadsheet users. It is well-established that spreadsheets can support a diverse range of possibilities for accomplishing a task [48]. Nardi and Miller's [85] earlier work suggested that this quality provides opportunities for mutual learning, because it enables collaboration between technical experts (e.g. a competent spreadsheet user) and non-technical domain experts (e.g. a business owner) when solving domain-relevant tasks. Our study, however, suggests that user collaborations are more sequential and distributed in nature. Users often alternate control and access over a spreadsheet, and there are unspoken expectations for when intervention from another user is considered acceptable. In some cases, only a single individual is given editing privileges (e.g. S1P1), while other users are limited to commenting only. Participants suggested that such turn-taking is necessary to avoid errors, miscommunication, and challenges around version control. These issues have been well-documented in previous literature [23, 91]. However, it also limits users' exposure to learning new and alternative practices from others.

A further observation from our study suggests that individual differences in spreadsheet expertise can exacerbate problems with collaboration. In particular, we note that while users embrace the flexibility of the spreadsheet, they also view their own spreadsheet use strategies to be highly personalised and 'non-standard,' though definitions of 'standards' are often vaguely defined. This perception makes users reluctant to engage in sharing practices, as seen with participants such as S2P10 (see section 4.2.2), where their spreadsheet was too specific to their own needs to be sent

to others. This desire to conform, or avoid highlighting one's 'other-ness,' seems to be linked to concerns around self-presentation. These concerns were most evident when participants described their social roles and any accompanying expectations of responsibility (when in positions of seniority e.g. S1P25) or expertise (due to length of their career e.g. S1P32). Earlier works have demonstrated that the social relationship between providers and recipients, along with concerns of organisational surveillance, are important factors affecting user acceptance of expertise sharing systems [40, 97]. Our work further emphasises how subjectivity and differences in user practices can exacerbate these social concerns in knowledge sharing around software.

Furthermore, providers who are concerned about their 'non-standard' spreadsheet practices will expend more effort to extract, filter, or condense their thought processes in a way which they believe would meet the needs of their recipients. This finding concurs with recent studies highlighting that authors are highly motivated to annotate and modify spreadsheets to facilitate collaboration [24, 108]. Our work expands these studies by emphasising that user motivations are as much about conformity as they are about communication, and that this nuance limits user willingness to share. Smith et al. [107] had previously observed that experienced users may refrain from using advanced spreadsheet features in collaborative contexts, and we have found further evidence to corroborate this finding (e.g. S1P21).

Overall, our study highlights how social concerns and motivations are deeply embedded in perceptions of spreadsheet interactions, to the extent that they interfere with collaborative behaviours. While normative beliefs about standard practices and concerns of self-presentation may make spreadsheet users reluctant to reveal their personal approaches, these tendencies may be moderated if spreadsheet software provided greater ease for standardising and communicating such standards between users. Similarly, improving synchronous collaboration—for example, by offering support for tracking edits, user history, and version control—could promote further opportunities for knowledge sharing.

5.2 Feature-rich software and the challenge of knowledge self-efficacy

Within the knowledge management literature, knowledge self-efficacy (i.e. an individual's confidence in the value of their knowledge contribution) is an important pre-requisite for expertise sharing to take place [54, 59, 86]. As a theoretical construct, self-efficacy is informed by an individual's personal experiences of desirable or undesirable consequences, social comparison, vicarious observation of others, or verbal persuasion [7]. Furthermore, it is captured in several behavioural theories commonly applied to knowledge sharing, such as social cognitive theory [7] and TPB [2, 3]. Our study placed a particular focus on how interactions with the spreadsheet may produce experiences which heighten or diminish one's knowledge self-efficacy. In the context of spreadsheet software, knowledge self-efficacy specifically concerns how an individual self-reports their own spreadsheet expertise.

Our findings show that self-evaluations of spreadsheet expertise are often mixed. Most participants expressed confidence in their ability to apply spreadsheets to address work-related needs. Several participants remarked on the simplicity of spreadsheet use, and shared their belief that most would know how to use a spreadsheet. Yet many of the same individuals were also cautious to credit themselves as knowledgeable. We noted several instances where the participants portrayed the spreadsheet as containing an insurmountable number of features, and framed their own expertise as a small proportion of the overall total.

These findings mirror a core problem previously identified in software learning research [28]. That is, users often fail to adapt advanced features and practices of feature-rich software even if they have long-term experience with using it [28, 81, 83, 88]. Part of the challenge, as discussed in section 5.1, is that feature-rich software offer multiple ways to accomplish a task [83, 85]. However,

due to this multiplicity, the process of learning and deciding what novel features are relevant may become effortful according to previous studies [62]. This may explain why our participants tend to re-use existing solutions, rather than exploring novel features. Furthermore, analyses of corporate spreadsheets suggest that most users rely only on a small fraction of available formulas for the majority of tasks [49, 56]. This corresponds to our participants' self-deprecating descriptions of their own expertise.

The simultaneously conflicting remarks about the 'simplicity' of spreadsheet use could be explained by the design of spreadsheet interfaces. Cockburn et al. [28] had previously suggested that intuitive user interfaces, designed with initial learnability in mind [83], may unintentionally hinder users' discovery of advanced features and practices (e.g. hotkey shortcuts). While expert hotkey usage is somewhat more simplistic in nature compared to expert formula use in spreadsheets, some participants did indeed link the simplicity of spreadsheet use to its interface. Previous analyses also suggest that spreadsheet interfaces incorporate several attributes conducive for initial learning and adoption, such as the grid, the taskbar, and their immediate reactivity to user interactions [48, 77]. We do not necessarily argue that the current designs deter user discovery of more complex features (e.g. formulas, VBA). However, it is possible that alternative spreadsheet representations which provide greater visibility of underlying computational processes (see [102]) may encourage greater adoption of advanced features.

In summary, we argue that the conflicted notions of spreadsheet expertise and self-efficacy observed may be attributed to two core factors. First, users are aware of the inherent complexity of the spreadsheet, and negatively evaluate their own expertise if they feel they are not sufficiently familiar with all its capabilities. Second, users may be disincentivised to explore and adopt more complex features (e.g. formulas), perhaps due to the intuitive and accessible nature of the spreadsheet's main interface. We suggest that these findings highlight the possible tension between balancing designs for initial learnability and the needs that arise in long-term learning [101]. Such considerations around software learnability directly affects the way users develop their knowledge self-efficacy, and thus are meaningful to consider for supporting knowledge sharing. However, further research is needed to establish whether feedback and framing interventions on self-efficacy will be effective for encouraging software-related knowledge sharing (e.g. [54, 86]).

Finally, while the role of knowledge self-efficacy in sharing is important in theory [2, 7, 54, 86], we acknowledge that the participants did not explicitly mention this concept in their discussions. This is partly related to the fact that interview data is dependent on interviewees' introspection abilities. Though an empirical relationship between spreadsheet knowledge self-efficacy and sharing has been identified in a recent study [117], we believe that further validation of the connection is necessary.

5.3 Implications for design

To address the challenges identified in section 5.1 and section 5.2, we propose the following considerations for researchers and technology designers.

5.3.1 Supporting users in identifying relevant features. The large number of available features in spreadsheet software makes it challenging for participants to identify what is useful or relevant for their needs. We suggest that spreadsheet systems could address these issues by organising features based on their relevance to users' specific needs. This can reduce the amount of information users must consider when learning. Previous studies have explored adaptive interfaces that allow users to customise their preferred subset of features [74, 82], or tailor the taskbar according to users' preferences and use history automatically [19, 37]. Other studies have explored task-centric interfaces, such as 'Workflows,' which organises available features based on specific tasks [67].

Some systems also provide suggestions on how to carry out a task [38, 68]. These approaches may help users to improve their proficiency, to notice differences between user preferences more easily, and therefore to discover previously unfamiliar features relevant to their work.

An alternative approach to identifying features or software-related information is learning from other users in the same work domain. One source would be recommendations from colleagues or teammates [12, 40], or members of a wider user community [72, 80, 114]. Previous studies have shown that this approach to learning is helpful for users to align their ways of use with others' [40]. For spreadsheet users, standardisation and training are often severely lacking [29, 70, 94], even for users within the same organisation [107]. Sharing knowledge between colleagues could help individuals develop mutually agreed standards for use.

However, as we have previously discussed, designing systems to support expertise sharing requires consideration of two main aspects: the quality of information extracted for recipients' benefits, and the agency, control, and ownership of providers. As Giannisakis et al. [40] suggest, passive sharing demands less effort but often causes a loss of control over shared information [39, 69], while manual sharing and annotation demands more effort, but is desirable for users who are willing to support others. Future research that explores when and where an automated form of expertise sharing is most appropriate will be helpful. A careful consideration of how systems supporting such automation are introduced into the specific culture of an organisation is necessary to ensure user acceptability and effectiveness.

5.3.2 Improving clarity around user expertise. An alternative approach for praxisware designers is to be more transparent regarding the inherent complexity of feature-rich software. This approach stands in contrast with the prevailing trend, which tends to favour simple designs and interfaces [22, 28]. As Sarkar argues, users are both capable and willing to commit to long-term software learning when the benefits are made clear. Therefore, designers should acknowledge, rather than obscure, the challenges associated with such learning, and empower users throughout the learning process. In line with Sarkar's argument [101], our participants were quick to point out the necessity and value of the spreadsheet within their work despite its many established complexities, frustrations, and challenges [23, 91]. However, most participants struggled to articulate their technical understanding of the spreadsheet software or the standards of quality expected from others, which led to difficulties when self-evaluating expertise. Exploring methods that offer clearer feedback and concrete performance indicators could help users better assess their expertise. This may reduce the perceived social costs of sharing software-related information, while encouraging further learning.

An example of such a feedback intervention is 'Skill-o-meter' [76], a widget which provides visual feedback of the time taken to complete tasks for the user's chosen method (e.g. mouse-based interaction), and compares it to the time of using a recommended alternative (e.g. hotkeys). The use of Skill-o-meter was shown to increase the usage of hotkeys among users. A related intervention called 'Search Dashboard' applies a similar concept to web searches, allowing users to compare their own search history to 'expert' searcher profiles and improve the efficiency of their own searching behaviours [11]. Another related technique is labelling learning materials, such as tutorials and videos, according to varying levels of difficulty, which could also help users strengthen their understanding of their own capabilities [100].

Traditional metrics for expertise are often hierarchical or linear. This is thus challenging to apply in the spreadsheet context, where the most efficient method for a task may not necessarily be useful in another. For software used by diverse user groups, it may be more appropriate to acknowledge different manifestations of expertise, rather than assign arbitrary value judgements and categorise users as either 'low' or 'high' expertise. Grossman and Fitzmaurice [41] suggest using descriptive labels of expertise based on usage efficiency, familiarity with features, and relevance of knowledge

to describe patterns in users' knowledge and preferences. For example, they describe 'naïve' experts as users who lack knowledge of relevant alternative interaction methods but are highly familiar and efficient with a small subset of features, and 'isolated' experts as users who may be efficient with a wide variety of interaction methods, but only for their specific domain purposes [41]. This approach may help users to acknowledge the validity of their preferred practices, while providing clarity as to how other methods may help meet their needs. Designers could also delineate different spreadsheet use strategies based on the values important to users, such as rating or providing measurements of formulas based on readability [51] and weakness to errors. This provides users with practical criteria with which they could self-evaluate expertise rather than arbitrary and irrelevant impressions.

6 Limitations

We recognise that data analysis and interpretation under an interpretivist stance may possibly be subject to the researcher's personal experiences and values [13, 75]. To mitigate this, we took measures where possible. We engaged in the principles of reflexivity and introduced alternative perspectives [17] from a second interview dataset in data analysis. While the first author was ultimately responsible for the final interpretation and presentation of these findings, using data collected by different researchers (non-authors) and incorporating discussions with co-authors provided a more rounded perspective in analysis.

The other limitation was that we did not collect demographic data concerning participants' cultures, a potential factor which can influence expertise sharing practices [4]. However, since most participants were recruited in the United Kingdom, it may be assumed that we do not have sufficient sample size to identify potential cultural trends. This issue could be explored further in future research. While the interviews were conducted by us, which allowed us to validate participant engagement, we did not administer attention checks during the survey, which may limit the validity of responses.

7 Conclusion

We provide two contributions in this paper. The first contribution is our investigation into the social barriers around spreadsheet expertise sharing. Through a qualitative analysis of interviews with 31 spreadsheet users, we identify how users struggle to adapt personalised use strategies to subjective standards of quality and to navigate timings for sharing. We also find that concerns over potential collaboration errors, conflicted self-evaluations of spreadsheet expertise, and wider social norms devaluing spreadsheet-related knowledge, further disincentivise expertise sharing practices. We highlight how these challenges are inextricably linked to the flexibility and complexity of feature-rich software. Our second contribution is a set of design recommendations based on our findings which may help address the challenges identified. We suggest that different design choices can support users in developing realistic expectations for expertise by: being more transparent about the complexity of using a feature-rich software, using taskbars which support personalisation and prioritise feature relevance, and emphasising descriptive labels of expertise over hierarchical labels. The findings of this paper suggest that considerations for alternative design choices may lead to different user dynamics and engagement in collaborative behaviours in the long term. By highlighting the sociotechnical nature of expertise sharing practices, our paper provides insights and reflections which may be useful not only for spreadsheet users, but could also inform future software design considerations.

Acknowledgments

References

[1] M. S. Ackerman and T. W. Malone. 1990. Answer Garden: A Tool for Growing Organizational Memory. *ACM SIGOIS Bulletin* 11, 2-3 (April 1990), 31–39. doi:10.1145/91478.91485

[2] Icek Ajzen. 1985. From Intentions to Actions: A Theory of Planned Behavior. In *Action Control*, Julius Kuhl and Jürgen Beckmann (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 11–39. doi:10.1007/978-3-642-69746-3_2

[3] Icek Ajzen. 1991. The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes* 50, 2 (Dec. 1991), 179–211. doi:10.1016/0749-5978(91)90020-T

[4] Alexandre Ardichvili, Martin Maurer, Wei Li, Tim Wentling, and Reed Stuedemann. 2006. Cultural Influences on Knowledge Sharing through Online Communities of Practice. *Journal of Knowledge Management* 10, 1 (Jan. 2006), 94–107. doi:10.1108/13673270610650139

[5] Alexander Ardichvili, Vaughn Page, and Tim Wentling. 2003. Motivation and Barriers to Participation in Virtual Knowledge-sharing Communities of Practice. *Journal of Knowledge Management* 7, 1 (March 2003), 64–77. doi:10.1108/13673270310463626

[6] Kenneth R. Baker, Lynn Foster-Johnson, Barry Lawson, and Stephen G. Powell. 2006. A Survey of MBA Spreadsheet Users. (2006).

[7] Albert Bandura. 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall, Englewood Cliffs, NJ.

[8] Albert Bandura. 2001. Social Cognitive Theory: An Agentic Perspective. *Annual Review of Psychology* 52, 1 (Feb. 2001), 1–26. doi:10.1146/annurev.psych.52.1.1

[9] Lyn Bartram, Michael Correll, and Melanie Tory. 2022. Untidy Data: The Unreasonable Effectiveness of Tables. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (Jan. 2022), 686–696. doi:10.1109/TVCG.2021.3114830

[10] Richard Baskerville and Alina Dulipovici. 2006. The Theoretical Foundations of Knowledge Management. *Knowledge Management Research & Practice* 4, 2 (May 2006), 83–105. doi:10.1057/palgrave.kmrp.8500090

[11] Scott Bateman, Jaime Teevan, and Ryen W. White. 2012. The Search Dashboard: How Reflection and Comparison Impact Search Behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin Texas USA, 1785–1794. doi:10.1145/2207676.2208311

[12] Scott S. Bateman, Carl A. Gutwin, and Gordon I. McCalla. 2013. Social Navigation for Loosely-Coupled Information Seeking in Tightly-Knit Groups Using Webwear. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. ACM, San Antonio Texas USA, 955–966. doi:10.1145/2441776.2441885

[13] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. *Qualitative HCI Research: Going Behind the Scenes*. Springer International Publishing, Cham. doi:10.1007/978-3-031-02217-3

[14] Gee W. Bock and Young-Gul Kim. 2002. Breaking the Myths of Rewards: An Exploratory Study of Attitudes about Knowledge Sharing. *Information Resources Management Journal* 15, 2 (April 2002), 14–21. doi:10.4018/irmj.2002040102

[15] Bock, Gee-Woo, Zmud, Robert W., Kim, Young-Gul, and Lee, Jae-Nam. 2005. Behavioral Intention Formation in Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organizational Climate. *MIS Quarterly* 29, 1 (2005), 87. jstor:10.2307/25148669 doi:10.2307/25148669

[16] Mark C. Bolino. 1999. Citizenship and Impression Management: Good Soldiers or Good Actors? *The Academy of Management Review* 24, 1 (Jan. 1999), 82. jstor:259038 doi:10.2307/259038

[17] Virginia Braun and Victoria Clarke. 2019. Reflecting on Reflexive Thematic Analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (Aug. 2019), 589–597. doi:10.1080/2159676X.2019.1628806

[18] Wayne Brekhus. 1998. A Sociology of the Unmarked: Redirecting Our Focus. *Sociological Theory* 16, 1 (1998), 34–51. jstor:202213

[19] Andrea Bunt, Cristina Conati, and Joanna McGrenere. 2007. Supporting Interface Customization Using a Mixed-Initiative Approach. In *Proceedings of the 12th International Conference on Intelligent User Interfaces*. ACM, Honolulu Hawaii USA, 92–101. doi:10.1145/1216295.1216317

[20] Andrea Bunt, Patrick Dubois, Ben Lafreniere, Michael A. Terry, and David T. Cormack. 2014. TaggedComments: Promoting and Integrating User Comments in Online Application Tutorials. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 4037–4046. doi:10.1145/2556288.2557118

[21] André Calero Valdez, Anne Kathrin Schaar, and Martina Ziefle. 2013. Personality Influences on Etiquette Requirements for Social Media in the Work Context. In *Human Factors in Computing and Informatics*, David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi, Gerhard Weikum, Andreas Holzinger, Martina Ziefle, Martin Hitz, and Matjaž Debevc (Eds.). Vol. 7946. Springer Berlin Heidelberg, Berlin, Heidelberg, 427–446. doi:10.1007/978-3-642-39062-3_27

[22] John M. Carroll and Mary Beth Rosson. 1987. Paradox of the Active User. In *Interfacing Thoughts: Cognitive Aspects of Human-Computer Interaction*. MIT Press, Cambridge, 80–111.

[23] Jonathan P Caulkins, Erika Layne Morrison, and Timothy Weidemann. 2007. Spreadsheet Errors and Decision Making: Evidence from Field Interviews. *Journal of Organisational and End User Computing* 19 (2007).

[24] George Chalhoub and Advait Sarkar. 2022. "It's Freedom to Put Things Where My Mind Wants": Understanding and Improving the User Experience of Structuring Data in Spreadsheets. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–24. doi:10.1145/3491102.3501833

[25] Yuexi Chen, Vlad I Morariu, Anh Truong, and Zhicheng Liu. 2024. TutoAI: A Cross-Domain Framework for AI-assisted Mixed-Media Tutorial Creation on Physical Tasks. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–17. doi:10.1145/3613904.3642443

[26] Parmit K. Chilana, Nathaniel Hudson, Srinjita Bhaduri, Prashant Shashikumar, and Shaun Kane. 2018. Supporting Remote Real-Time Expert Help: Opportunities and Challenges for Novice 3D Modelers. In *2018 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, Lisbon, 157–166. doi:10.1109/VLHCC.2018.8506568

[27] Luke Church, Emma Söderberg, and Martin Höst. 2023. My Space, Our Space, Their Space: A First Glance at Developers' Experience of Spaces. In *Companion Proceedings of the 7th International Conference on the Art, Science, and Engineering of Programming*. ACM, Tokyo Japan, 48–53. doi:10.1145/3594671.3594680

[28] Andy Cockburn, Carl Gutwin, Joey Scarr, and Sylvain Malacria. 2014. Supporting Novice to Expert Transitions in User Interfaces. *Comput. Surveys* 47, 2 (Nov. 2014), 31:1–31:36. doi:10.1145/2659796

[29] Paul B. Cragg and Malcolm King. 1993. Spreadsheet Modelling Abuse: An Opportunity for OR? *Journal of the Operational Research Society* 44, 8 (Aug. 1993), 743–752. doi:10.1057/jors.1993.133

[30] Grenville J. Croll. 2007. The Importance and Criticality of Spreadsheets in the City of London. (2007). doi:10.48550/ARXIV.0709.4063

[31] Jácome Cunha, Mihai Dan, Martin Erwig, Danila Fedorin, and Alex Grejuc. 2020. Explaining Spreadsheets with Spreadsheets (Short Paper). *ACM SIGPLAN Notices* 53, 9 (April 2020), 161–167. doi:10.1145/3393934.3278136

[32] Fernando Diaz, Bhaskar Mitra, Michael D. Ekstrand, Asia J. Biega, and Ben Carterette. 2020. Evaluating Stochastic Rankings with Expected Exposure. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. ACM, Virtual Event Ireland, 275–284. doi:10.1145/3340531.3411962

[33] Benjamin D. Douglas, Patrick J. Ewell, and Markus Brauer. 2023. Data Quality in Online Human-Subjects Research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLOS ONE* 18, 3 (March 2023), e0279720. doi:10.1371/journal.pone.0279720

[34] Patrick Dubois, Volodymyr Dziubak, and Andrea Bunt. 2017. Tell Me More! Soliciting Reader Contributions to Software Tutorials. *Proceedings of Graphics Interface 2017* Edmonton (2017), 8 pages, 632.60 KB. doi:10.20380/GI2017.03

[35] Muhammad Ashraf Fauzi. 2019. Knowledge Sharing in Asia Pacific via Virtual Community Platform: A Systematic Review. *International Journal of Web Based Communities* 15, 4 (2019), 368. doi:10.1504/IJWBC.2019.103186

[36] Muhammad Ashraf Fauzi, Mai Nguyen, and Ashish Malik. 2024. Knowledge Sharing and Theory of Planned Behavior: A Bibliometric Analysis. *Journal of Knowledge Management* 28, 2 (Jan. 2024), 293–311. doi:10.1108/JKM-11-2022-0933

[37] Leah Findlater, Karyn Moffatt, Joanna McGrenere, and Jessica Dawson. 2009. Ephemeral Adaptation: The Use of Gradual Onset to Improve Menu Selection Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Boston MA USA, 1655–1664. doi:10.1145/1518701.1518956

[38] C. Ailie Fraser, Mira Dontcheva, Holger Winnemöller, Sheryl Ehrlich, and Scott Klemmer. 2016. DiscoverySpace: Suggesting Actions in Complex Software. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, Brisbane QLD Australia, 1221–1232. doi:10.1145/2901790.2901849

[39] Anna Gausen, Bhaskar Mitra, and Siân Lindley. 2023. A Framework for Exploring the Consequences of AI-Mediated Enterprise Knowledge Access and Identifying Risks to Workers. arXiv:2312.10076 [cs]

[40] Emmanouil Giannisakis, Jessalyn Alvina, Andrea Bunt, Parmit Chilana, and Joanna McGrenere. 2022. Promoting Feature Awareness by Leveraging Collaborators' Usage Habits in Collaborative Editors. In *Graphics Interface*.

[41] Tovi Grossman and George Fitzmaurice. 2015. An Investigation of Metrics for the In Situ Detection of Software Expertise. *Human–Computer Interaction* 30, 1 (Jan. 2015), 64–102. doi:10.1080/07370024.2014.881668

[42] Tovi Grossman, George Fitzmaurice, and Ramtin Attar. 2009. A Survey of Software Learnability: Metrics, Methodologies and Guidelines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Boston MA USA, 649–658. doi:10.1145/1518701.1518803

[43] Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2010. Chronicle: Capture, Exploration, and Playback of Document Workflow Histories. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*. ACM, New York New York USA, 143–152. doi:10.1145/1866029.1866054

[44] European Spreadsheet Risks Interest Group. 2022. Horror Stories.

[45] Lisa A Guion, David C Diehl, and Debra McDonald. 2002. Triangulation: Establishing the Validity of Qualitative Studies. University of Florida Institute of Food and Agricultural Sciences working paper FCS6014.

[46] Stuart Hall. 1997. Representation: Cultural Representations and Signifying Practices. In *The Work of Representation*, Stuart Hall (Ed.). Sage, London, 13–74.

[47] D.G. Hendry and T.R.G. Green. 1993. CogMap: A Visual Description Language for Spreadsheets. *Journal of Visual Languages & Computing* 4, 1 (March 1993), 35–54. doi:10.1006/jvlc.1993.1003

[48] Felienne Hermans, Bas Jansen, Sohon Roy, Efthimia Aivaloglou, Alaaeddin Swidan, and David Hoepelman. 2016. Spreadsheets Are Code: An Overview of Software Engineering Approaches Applied to Spreadsheets. In *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*. IEEE, Suita, Osaka, Japan, 56–65. doi:10.1109/SANER.2016.86

[49] Felienne Hermans and Emerson Murphy-Hill. 2015. Enron's Spreadsheets and Related Emails: A Dataset and Analysis. In *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*. IEEE, Florence, Italy, 7–16. doi:10.1109/ICSE.2015.129

[50] Felienne Hermans, Martin Pinzger, and Arie Van Deursen. 2011. Supporting Professional Spreadsheet Users by Generating Leveled Dataflow Diagrams. In *Proceedings of the 33rd International Conference on Software Engineering*. ACM, Waikiki, Honolulu HI USA, 451–460. doi:10.1145/1985793.1985855

[51] Felienne Hermans, Martin Pinzger, and Arie Van Deursen. 2012. Detecting Code Smells in Spreadsheet Formulas. In *2012 28th IEEE International Conference on Software Maintenance (ICSM)*. IEEE, Trento, Italy, 409–418. doi:10.1109/ICSM.2012.6405300

[52] Chin-Lung Hsu and Judy Chuan-Chuan Lin. 2008. Acceptance of Blog Usage: The Roles of Technology Acceptance, Social Influence and Knowledge Sharing Motivation. *Information & Management* 45, 1 (Jan. 2008), 65–74. doi:10.1016/j.im.2007.11.001

[53] Meng-Hsiang Hsu, Teresa L. Ju, Chia-Hui Yen, and Chun-Ming Chang. 2007. Knowledge Sharing Behavior in Virtual Communities: The Relationship between Trust, Self-Efficacy, and Outcome Expectations. *International Journal of Human-Computer Studies* 65, 2 (Feb. 2007), 153–169. doi:10.1016/j.ijhcs.2006.09.003

[54] Shin-Yuan Hung, Hui-Min Lai, and Yu-Che Chou. 2015. Knowledge-Sharing Intention in Professional Virtual Communities: A Comparison between Posters and Lurkers. *Journal of the Association for Information Science and Technology* 66, 12 (2015), 2494–2510. doi:10.1002/asi.23339

[55] Omayma Husain, Naomie Salim, Rose Alinda Alias, Samah Abdelsalam, and Alzubair Hassan. 2019. Expert Finding Systems: A Systematic Review. *Applied Sciences* 9, 20 (Oct. 2019), 4250. doi:10.3390/app9204250

[56] Bas Jansen. 2015. Enron versus EUSES: A Comparison of Two Spreadsheet Corpora. arXiv:1503.04055 [cs] doi:10.48550/arXiv.1503.04055

[57] Murray E Jennex. 2006. Knowledge Management in Modern Organizations. (2006).

[58] Nikhita Joshi, Justin Matejka, Fraser Anderson, Tovi Grossman, and George Fitzmaurice. 2020. MicroMentor: Peer-to-Peer Software Help Sessions in Three Minutes or Less. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. doi:10.1145/3313831.3376230

[59] Atreyi Kankanhalli, Bernard C. Y. Tan, and Kwok-Kee Wei. 2005. Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation. *MIS Quarterly* 29, 1 (2005), 113–143. jstor:25148670 doi:10.2307/25148670

[60] Bennett Kankuzi and Jorma Sajaniemi. 2016. A Mental Model Perspective for Tool Development and Paradigm Shift in Spreadsheets. *International Journal of Human-Computer Studies* 86 (Feb. 2016), 149–163. doi:10.1016/j.ijhcs.2015.10.005

[61] Kimia Kiani, Parmit K. Chilana, Andrea Bunt, Tovi Grossman, and George Fitzmaurice. 2020. "I Would Just Ask Someone": Learning Feature-Rich Design Software in the Modern Workplace. In *2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, Dunedin, New Zealand, 1–10. doi:10.1109/VL/HCC50065.2020.9127288

[62] Kimia Kiani, George Cui, Andrea Bunt, Joanna McGrenere, and Parmit K. Chilana. 2019. Beyond "One-Size-Fits-All": Understanding the Diversity in How Software Newcomers Discover and Make Use of Help Resources. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–14. doi:10.1145/3290605.3300570

[63] Nigel King. 1998. Template Analysis. In *Qualitative Methods and Analysis in Organizational Research: A Practical Guide*. Sage Publications Ltd, Thousand Oaks, CA, 118–134.

[64] Andrea Kohlhase, Michael Kohlhase, and Ana Guseva. 2015. Context in Spreadsheet Comprehension. (2015), 21–27.

[65] Benjamin Lafreniere, Andrea Bunt, Matthew Lount, Filip Krynicki, and Michael A. Terry. 2011. AdaptableGIMP: Designing a Socially-Adaptable Interface. In *Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface*

Software and Technology (UIST '11 Adjunct). Association for Computing Machinery, New York, NY, USA, 89–90. doi:10.1145/2046396.2046437

[66] Ben Lafreniere, Andrea Bunt, Matthew Lount, and Michael Terry. 2013. Understanding the Roles and Uses of Web Tutorials. *Proceedings of the International AAAI Conference on Web and Social Media* 7, 1 (2013), 303–310. doi:10.1609/icwsm.v7i1.14413

[67] Benjamin Lafreniere, Andrea Bunt, and Michael Terry. 2014. Task-Centric Interfaces for Feature-Rich Software. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*. ACM, Sydney New South Wales Australia, 49–58. doi:10.1145/2686612.2686620

[68] Benjamin Lafreniere, Parmit K. Chilana, Adam Fourney, and Michael A. Terry. 2015. *These Aren't the Commands You're Looking For*: Addressing False Feedforward in Feature-Rich Software. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, Charlotte NC USA, 619–628. doi:10.1145/2807442.2807482

[69] Ida Larsen-Ledet, Bhaskar Mitra, and Siân Lindley. 2022. Ethical and Social Considerations in Automatic Expert Identification and People Recommendation in Organizational Knowledge Management Systems. arXiv:2209.03819 [cs]

[70] Barry R. Lawson, Kenneth R. Baker, Stephen G. Powell, and Lynn Foster-Johnson. 2009. A Comparison of Spreadsheet Users with Different Levels of Experience. *Omega* 37, 3 (June 2009), 579–590. doi:10.1016/j.omega.2007.12.004

[71] Soila Lemmetty and Kaija Collin. 2020. Self-Directed Learning as a Practice of Workplace Learning: Interpretative Repertoires of Self-Directed Learning in ICT Work. *Vocations and Learning* 13, 1 (April 2020), 47–70. doi:10.1007/s12186-019-09228-x

[72] Wei Li, Justin Matejka, Tovi Grossman, Joseph A. Konstan, and George Fitzmaurice. 2011. Design and Evaluation of a Command Recommendation System for Software Applications. *ACM Transactions on Computer-Human Interaction* 18, 2 (June 2011), 1–35. doi:10.1145/1970378.1970380

[73] Xiaolin Lin and Xuequn Wang. 2020. Examining Gender Differences in People's Information-Sharing Decisions on Social Networking Sites. *International Journal of Information Management* 50 (Feb. 2020), 45–56. doi:10.1016/j.ijinfomgt.2019.05.004

[74] Wendy E. Mackay. 1991. Triggers and Barriers to Customizing Software. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Reaching through Technology - CHI '91*. ACM Press, New Orleans, Louisiana, United States, 153–160. doi:10.1145/108844.108867

[75] Anna Madill, Abbie Jordan, and Caroline Shirley. 2000. Objectivity and Reliability in Qualitative Analysis: Realist, Contextualist and Radical Constructionist Epistemologies. *British Journal of Psychology* 91, 1 (Feb. 2000), 1–20. doi:10.1348/000712600161646

[76] Sylvain Malacria, Joey Scarr, Andy Cockburn, Carl Gutwin, and Tovi Grossman. 2013. Skillometers: Reflective Widgets That Motivate and Help Users to Improve Performance. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*. ACM, St. Andrews Scotland, United Kingdom, 321–330. doi:10.1145/2501988.2501996

[77] John H. Maloney and Randall B. Smith. 1995. Directness and Liveness in the Morphic User Interface Construction Environment. In *Proceedings of the 8th Annual ACM Symposium on User Interface and Software Technology*. ACM, Pittsburgh Pennsylvania USA, 21–28. doi:10.1145/215585.215636

[78] Maria Mårtensson. 2000. A Critical Review of Knowledge Management as a Management Tool. *Journal of Knowledge Management* 4, 3 (Sept. 2000), 204–216. doi:10.1108/13673270010350002

[79] Justin Matejka, Tovi Grossman, and George Fitzmaurice. 2013. Patina: Dynamic Heatmaps for Visualizing Application Usage. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Paris France, 3227–3236. doi:10.1145/2470654.2466442

[80] Justin Matejka, Wei Li, Tovi Grossman, and George Fitzmaurice. 2009. CommunityCommands: Command Recommendations for Software Applications. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*. ACM, Victoria BC Canada, 193–202. doi:10.1145/1622176.1622214

[81] Joanna McGrenere, Ronald M. Baecker, and Kellogg S. Booth. 2002. An Evaluation of a Multiple Interface Design Solution for Bloated Software. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Minneapolis Minnesota USA, 164–170. doi:10.1145/503376.503406

[82] Joanna McGrenere, Ronald M. Baecker, and Kellogg S. Booth. 2007. A Field Evaluation of an Adaptable Two-Interface Design for Feature-Rich Software. *ACM Transactions on Computer-Human Interaction* 14, 1 (May 2007), 3. doi:10.1145/1229855.1229858

[83] Joanna Mcgrenere and Gale Moore. 2000. Are We All In the Same "Bloat"?. In *Proceedings of the Graphics Interface 2000 Conference*. 187–196.

[84] Henry Mintzberg. 1980. Structure in 5's: A Synthesis of the Research on Organization Design. *Management Science* 26, 3 (March 1980), 322–341. doi:10.1287/mnsc.26.3.322

[85] Bonnie A. Nardi and James R. Miller. 1991. Twinkling Lights and Nested Loops: Distributed Problem Solving and Spreadsheet Development. *International Journal of Man-Machine Studies* 34, 2 (Feb. 1991), 161–184. doi:10.1016/0020-7373(91)90040-E

[86] Tuyet-Mai Nguyen, Liem Viet Ngo, and Gary Gregory. 2020. Motivation in Organisational Online Knowledge Sharing. *Journal of Knowledge Management* 26, 1 (May 2020), 102–125. doi:10.1108/JKM-09-2020-0664

[87] Tuyet-Mai Nguyen, Tuan Phong Nham, Fabian Jintae Froese, and Ashish Malik. 2019. Motivation and Knowledge Sharing: A Meta-Analysis of Main and Moderating Effects. *Journal of Knowledge Management* 23, 5 (June 2019), 998–1016. doi:10.1108/JKM-01-2019-0029

[88] Erik Nilsen, HeeSen Jong, Judith S. Olson, Kevin Biolsi, Henry Rueter, and Sharon Mutter. 1993. The Growth of Software Skill: A Longitudinal Look at Learning & Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '93.* ACM Press, Amsterdam, The Netherlands, 149–156. doi:10.1145/169059.169126

[89] Ikujirō Nonaka and Hirotaka Takeuchi. 1995. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation.* Oxford University Press, New York.

[90] Midas Nouwens and Clemens Nylandsted Klokmose. 2018. The Application and Its Consequences for Non-Standard Knowledge Work. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* ACM, Montreal QC Canada, 1–12. doi:10.1145/3173574.3173973

[91] Ray Panko. 2016. What We Don't Know About Spreadsheet Errors Today: The Facts, Why We Don't Believe Them, and What We Need to Do. (2016). doi:10.48550/ARXIV.1602.02601

[92] Sharoda A. Paul. 2016. Find an Expert: Designing Expert Selection Interfaces for Formal Help-Giving. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16).* Association for Computing Machinery, New York, NY, USA, 3038–3048. doi:10.1145/2858036.2858131

[93] Eyal Peer, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer. 2022. Data Quality of Platforms and Panels for Online Behavioral Research. *Behavior Research Methods* 54, 4 (Aug. 2022), 1643–1662. doi:10.3758/s13428-021-01694-3

[94] J.D. Pemberton and A.J. Robson. 2000. Spreadsheets in Business. *Industrial Management & Data Systems* 100, 8 (Nov. 2000), 379–388. doi:10.1108/02635570010353938

[95] David Pilgrim. 2014. Some Implications of Critical Realism for Mental Health Research. *Social Theory & Health* 12, 1 (Feb. 2014), 1–21. doi:10.1057/sth.2013.17

[96] Volkmar Pipek and Volker Wulf. 2003. Pruning the Answer Garden: Knowledge Sharing in Maintenance Engineering. In *ECSCW 2003*, Kari Kuutti, Eija Helena Karsten, Geraldine Fitzpatrick, Paul Dourish, and Kjeld Schmidt (Eds.). Springer Netherlands, Dordrecht, 1–20. doi:10.1007/978-94-010-0068-0_1

[97] Volkmar Pipek, Volker Wulf, and Aditya Johri. 2012. Bridging Artifacts and Actors: Expertise Sharing in Organizational Ecosystems. *Computer Supported Cooperative Work (CSCW)* 21, 2-3 (June 2012), 261–282. doi:10.1007/s10606-011-9138-y

[98] Fatemeh Riahi, Zainab Zolaktaf, Mahdi Shafiei, and Evangelos Milios. 2012. Finding Expert Users in Community Question Answering. In *Proceedings of the 21st International Conference on World Wide Web (WWW '12 Companion).* Association for Computing Machinery, New York, NY, USA, 791–798. doi:10.1145/2187980.2188202

[99] Andreas Riege. 2005. Three-dozen Knowledge-sharing Barriers Managers Must Consider. *Journal of Knowledge Management* 9, 3 (June 2005), 18–35. doi:10.1108/13673270510602746

[100] Shahed Anzarus Sabab, Adnan Khan, Parmit K. Chilana, Joanna McGrenere, and Andrea Bunt. 2020. An Automated Approach to Assessing an Application Tutorial's Difficulty. In *2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC).* IEEE, Dunedin, New Zealand, 1–10. doi:10.1109/VL/HCC50065.2020.9127271

[101] Advait Sarkar. 2023. Should Computers Be Easy To Use? Questioning the Doctrine of Simplicity in User Interface Design. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23).* Association for Computing Machinery, New York, NY, USA, 1–10. doi:10.1145/3544549.3582741

[102] Advait Sarkar and Andrew D Gordon. 2018. How Do People Learn to Use Spreadsheets? (Work in Progress). In *Proceedings of the 29th Annual Conference of the Psychology of Programming Interest Group (PPIG 2018).*

[103] Advait Sarkar, Sruti Srinivasa Ragavan, Jack Williams, and Andrew D. Gordon. 2022. End-User Encounters with Lambda Abstraction in Spreadsheets: Apollo's Bow or Achilles' Heel?. In *2022 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC).* IEEE, Roma, Italy, 1–11. doi:10.1109/VL/HCC53370.2022.9833131

[104] C. Scaffidi, M Shaw, and B Myers. 2005. Estimating the Numbers of End Users and End User Programmers | IEEE Conference Publication | IEEE Xplore. In *2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05).* 207–214. doi:10.1109/VLHCC.2005.34

[105] Thomas Schoegje, Lynda Hardman, Arjen De Vries, and Toine Pieters. 2024. Improving Expert Search Effectiveness: Comparing Ways to Rank and Present Search Results. In *Proceedings of the 2024 ACM SIGIR Conference on Human Information Interaction and Retrieval.* ACM, Sheffield United Kingdom, 56–65. doi:10.1145/3627508.3638296

[106] (Spreadsheet Engineering Research Project) SERP. 2005. SERP Survey Results. https://faculty.tuck.dartmouth.edu/images/uploads/faculty/serp/serp_results.pdf.

[107] Justin Smith, Justin A. Middleton, and Nicholas A. Kraft. 2017. Spreadsheet Practices and Challenges in a Large Multinational Conglomerate. In *2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC).*

155–163. doi:10.1109/VLHCC.2017.8103463

[108] Sruti Srinivasa Ragavan, Advait Sarkar, and Andrew D Gordon. 2021. Spreadsheet Comprehension: Guesswork, Giving Up and Going Back to the Author. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–21. doi:10.1145/3411764.3445634

[109] Gabriel Szulanski. 1996. Exploring Internal Stickiness: Impediments to the Transfer of Best Practice within the Firm. *Strategic Management Journal* 17, S2 (1996), 27–43. doi:10.1002/smj.4250171105

[110] Asha Thomas and Vikas Gupta. 2021. The Role of Motivation Theories in Knowledge Sharing: An Integrative Theoretical Reviews and Future Research Agenda. *Kybernetes* 51, 1 (Jan. 2021), 116–140. doi:10.1108/K-07-2020-0465

[111] Michael B. Twidale. 2005. Over the Shoulder Learning: Supporting Brief Informal Learning. *Computer Supported Cooperative Work (CSCW)* 14, 6 (Dec. 2005), 505–547. doi:10.1007/s10606-005-9007-7

[112] Nan Wang, Jielin Yin, Zhenzhong Ma, and Maolin Liao. 2022. The Influence Mechanism of Rewards on Knowledge Sharing Behaviors in Virtual Communities. *Journal of Knowledge Management* 26, 3 (March 2022), 485–505. doi:10.1108/JKM-07-2020-0530

[113] Sheng Wang and Raymond A. Noe. 2010. Knowledge Sharing: A Review and Directions for Future Research. *Human Resource Management Review* 20, 2 (June 2010), 115–131. doi:10.1016/j.hrmr.2009.10.001

[114] Xu Wang, Benjamin Lafreniere, and Tovi Grossman. 2018. Leveraging Community-Generated Videos and Command Logs to Classify and Recommend Software Workflows. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–13. doi:10.1145/3173574.3173859

[115] Annick Willem and Harry Scarbrough. 2006. Social Capital and Political Bias in Knowledge Sharing: An Exploratory Study. *Human Relations* 59, 10 (Oct. 2006), 1343–1370. doi:10.1177/0018726706071527

[116] Haolun Wu, Bhaskar Mitra, Chen Ma, Fernando Diaz, and Xue Liu. 2022. Joint Multisided Exposure Fairness for Recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, Madrid Spain, 703–714. doi:10.1145/3477495.3532007

[117] Qing Nancy Xia, Advait Sarkar, Duncan P. Brumby, and Anna Cox. 2024. The Paradox of Spreadsheet Self-Efficacy: Social Incentives for Informal Knowledge Sharing in End-User Programming. In *2024 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. 81–88. doi:10.1109/VL/HCC60511.2024.00019

[118] Chi-Lan Yang, Chien Wen (Tina) Yuan, and Hao-Chuan Wang. 2019. When Knowledge Network Is Social Network: Understanding Collaborative Knowledge Transfer in Workplace. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–23. doi:10.1145/3359266

[119] Pingjing Yang, Ti-Chung Cheng, Sajjadur Rahman, Mangesh Bendre, Karrie Karahalios, and Aditya Parameswaran. 2020. Understanding Data Analysis Workflows on Spreadsheets: Roadblocks and Opportunities. (2020).

[120] Li Zhao and Brian Detlor. 2023. Towards a Contingency Model of Knowledge Sharing: Interaction between Social Capital and Social Exchange Theories. *Knowledge Management Research & Practice* 21, 1 (Jan. 2023), 197–209. doi:10.1080/14778238.2020.1866444

[121] Li Zhao, Brian Detlor, and Catherine E. Connelly. 2016. Sharing Knowledge in Social Q&A Sites: The Unintended Consequences of Extrinsic Motivation. *Journal of Management Information Systems* 33, 1 (Jan. 2016), 70–100. doi:10.1080/07421222.2016.1172459

A Ethics and Data Protection

In both studies, participants were fully informed of the procedures of the interview, the types of audio and visual data that would be collected, and the confidential nature of the personal information collected. All participants filled out a consent form prior to data collection. All participants, including those who filled out the survey only in Study 1, took part on a voluntary basis and were reimbursed for their time.

For Study 1, data protection authorisation and ethics clearance were provided by the university in accordance with regulations in the host institution. Participants were guided on how to censor or anonymise data in their spreadsheets to protect potentially sensitive or confidential data in the spreadsheet walkthrough. Audio and video recordings were made on Microsoft Teams and stored in a secure, institutional cloud accessible only to the researchers. Spreadsheet sharing was only done via screen-sharing; we did not request copies of participants' spreadsheets, nor did we have direct access to the spreadsheets and their data at any point.

Study 2, which was conducted by the second author, had been cleared through a previous institutional ethics review at the institution of that author. The study protocol, recruitment strategy,

and consent forms were reviewed and approved. Over the course of analysing results for Study 1, when it became apparent that the interview data collected in Study 2 could be relevant in supporting the final analysis, a further institutional ethics review was conducted to grant appropriate consent and ethical approval, which involved further anonymisation and redaction of the Study 2 transcripts, to allow for data sharing with the first author, who had not been involved in the original data collection phase of Study 2.