DOI: 10.1002/pra2.270

LONG PAPERS

WILEY

Opening up the black box of scholarly synthesis: Intermediate products, processes, and tools

Xin Qian | **Katrina Fenlon** | **Wayne Lutters** | **Joel Chan**

College of Information Studies, University of Maryland, College Park, Maryland, USA

Correspondence
Xin Qian, College of Information Studies, University of Maryland, College Park, MD 20740.
Email: xinq@umd.edu

Abstract

Synthesis is a foundational scholarly product that generates new conceptual wholes from independent intellectual sources. But effective synthesis products—such as literature reviews—are rare, in part due to inadequate support from existing tools and information systems. A detailed, situated understanding of the work practices behind synthesis is necessary to inform the development of synthesis tools. Previous work in scholarly primitives, active reading, and sensemaking provide partial explanations of aspects of synthesis, but a detailed explanation of scholarly synthesis, specifically, is lacking. This paper presents a foundational empirical examination of the work practices behind synthesis to address the gap, focusing on unpacking the intermediate products, processes, and tools through in-depth contextual interviews with scholars. Results shed light on the distinctive intermediate products generated during synthesis—including in-source annotations, per-source summaries, and cross-source syntheses—as well as effortful processes for nonlinear progression of these intermediate products towards a final synthesis product. These products and practices were also embedded in a complex ecology of creative reappropriated tools. This work enriches understanding of the complex scholarly practices that produce synthesis and opens up a research agenda for understanding and supporting scholarly synthesis more generally.

KEYWORDS

active reading, interdisciplinary research, literature review, scholarly primitive, sensemaking

1 | INTRODUCTION

In 2011, an economist traced the genesis of her fruitful research agenda to a "masterful survey of the literature on health, education, labor markets, and household behavior in development economics...[that] made two things very apparent: the questions could not be more important, and the answers were mostly unsatisfactory" (p.5). The economist was Esther Duflo, who shared the Nobel Prize in Economics in 2019[1] for this work on experimental approaches to alleviating global poverty.

This example illustrates how **synthesis**—the creation of a new conceptual from independent intellectual sources—is a foundational process in scholarly work. The products of synthesis may take familiar forms like a literature review or systematic review. As Strike and Posner (1983) note, effective synthesis products can clarify inconsistencies and resolve tensions over multiple sources,

83rd Annual Meeting of the Association for Information Science & Technology October 25-29, 2020.

and enable progressive problem shifts with increased explanatory and predictive capacity, expand empirical evidence and scope of application, and strengthen theoretical support. In short, an effective synthesis is generative: it enables innovation in scholarly and scientific work.

Despite their importance, effective synthesis products are rare. While most published scholarship includes some review of prior literature, a minority truly synthesize prior knowledge. Case in point: two empirically grounded studies of committee review comments on the literature review sections of doctoral dissertations demonstrate that the bar for synthesis is substantially lower in practice than faculty's stated expectations (Denholm & Philpott, 2009; Holbrook, Bourke, Lovat, & Dally, 2004). For example, graduate students in psychology who are doing synthesis often use a sequential flow to organize related literature, instead of generating a new conceptual schema to integrate the prior literature (Froese, Gantz, & Henry, 1998; Granello, 2001). The problem is not confined to beginning researchers: as Bem (1995) notes, "authors of literature reviews are at risk for producing mind-numbing lists of citations and findings that resemble a phone book—impressive case, lots of numbers, but not much plot. In contrast, a coherent review emerges only from a coherent conceptual structuring of the topic itself." (p. 172). Frustrated by this tendency, some editors have developed guides for how to write reviews that synthesize prior work to enable new conceptual breakthroughs (Bem, 1995; Webster & Watson, 2002).

Effective synthesis is rare because it is arduous and inadequately supported by existing tools and information systems. Designing tools and information systems for synthesis requires navigating the considerable variation among individual researchers' practices and the nature of their research products. These practices and products are situated within a web of sociotechnical factors—including the domain of research, disciplinary traditions, team collaboration, research resources, tool dependencies—which affect successful synthesis in unknown ways.

Consider an illustrative scenario: Amani and Han are new scholars in Dr. Kopek's interdisciplinary health science lab focused on designing mobile games for behavior change. Both are working to identify the boundaries of a knowledge space. Amani is searching for models of effective behavior change across the social sciences, a process she describes as simultaneously an overwhelming flood of possibilities and a continual defeat of dead-ends. While she maintains a Zotero (Puckett, 2011) folder with citations and synopses to share with Dr. Kopek and the team, she also maintains for herself a hand-sketched map of the different concepts and theories and their promising interconnections. She is currently working on a white paper that distills her map and relates it to particular aspects of the research problem.

Han, on the other hand, is exploring game mechanics and their impact on players' self-efficacy. While there is a growing academic literature on this, he finds that the most significant insights come from information leaked from proprietary industry reports of playtesting, which he finds on blogs. He maintains a running, loosely structured Excel spreadsheet with pairs of "mechanic" and "impact" claims, the source, and initial gut reaction. He also keeps a folder of screenshots from Twitch gameplay streams as personal memory triggers of examples of these mechanics at work. He wants to compile a video of clips demonstrating the five highest impact designs, with links to research explaining why those mechanisms might be successful and to share out with the team. In both cases, Amani and Han are engaged in more than a traditional systematic literature review. Instead, they are seeking to synthesize what is known to motivate new understanding. However, they rely on divergent workflows and tools and produce fundamentally different kinds of outcomes along the way.

To develop the next generation of tools for synthesis work, we need a detailed, situated understanding of the **work practices** behind synthesis products. This effort should address critical questions about the fundamental "data structures" and "operations" in synthesis processes. Answers should encompass both the *intermediate products*, as well as the *processes* that produce and operate on them, moving from the raw input of experience and literature to the valuable final product of a synthesis. Informed by prior ethnomethodological insight (Cetina, 1999; Orr, 1996; Suchman, 1987), our explanatory power and resultant ability to design tools will be significantly enhanced by treating the particulars of where/how the practices and products are *situated* in the world. Such an understanding would allow us to interrogate the key affordances and cost structures involved in the practice of synthesis, and how they might vary across the tools, as well as identify the most critical pain points and opportunities for augmentation and optimization. In short, we need to open up the black box between **sources** and **synthesis**.

There is a rich scholarly tradition within Information Science that examines the physical, social, and cognitive processes and products of research and learning. A long strand of research on information seeking and information practices has sought to distill descriptive studies of scholarly processes into sets of *primitives*, or fundamental information activities, that are common across disciplines, such as "browsing," "chaining," "accessing," "notetaking," and "re-reading." (Blanke & Hedges, 2013; Ellis, 1993; Palmer, Teffeau, & Pirmann, 2009; Unsworth, 2000; Vilar, 2015). In their synthesis of the literature on primitives, Palmer et al. (2009) show how 20 discrete primitives function across disciplines as building-blocks for higher-level scholarly activities such as searching, collecting, reading,

writing, and collaborating. Building on this line of work, recent efforts to formally model workflows and scholarly communication as ontologies, for example, Gradmann et al. (2015) provide more holistic accounts of scholarly practices. They represent primitives as components within complex systems of actors, social contexts, domains, methods, goals, and tools. These models inform the design and development of research systems and tools to support widespread or cross-cutting research requirements; however, they do not explicitly address *synthesis* as a distinctive and problematic process.

Another window into the human practice of synthesis is through their material interaction with relevant documents. "Active reading" is a strand of research that focuses on understanding how people process individual papers. This work has revealed that people create intermediate products while reading, such as annotations and marginalia, and extracted content like snippets or quotes (Tashman & Edwards, 2011b). It also identified key subprocesses such as navigating between a paper's various subsections or arranging materials on a surface to gain an overview of their contents. However, this work does not explicitly situate these activites in the specific task of producing a scholarly synthesis. Instead, the highest level of processing studied is often the comparative clustering of multiple documents (O'Hara, 1996; O'hara & Sellen, 1997; Morris, Brush, & Meyers, 2007; Tashman & Edwards, 2011b, Tashman & Edwards, 2011a), stopping short of directly studying the intricate integration work in synthesis.

Models of sensemaking provide additional explanatory power for understanding synthesis, with several theories that specify objects and operations during sensemaking (for a recent comprehensive review of sensemaking models, see (Zhang & Dagobert, 2014)). These theories might help characterize the products and processes of synthesis. For example, the "Learning Loop Complex" models sensemaking as cycles through three processes: searching for good *representations* to capture salient information from the *data*, instantiating representations (originally termed as "*encodons*") with captured information, and shifting the instantiated representations to incorporate missing but relevant *data*, to fit task operations (Russell, Stefik, Pirolli, & Card, 1993). The "Data/Frame Model" models sensemaking as cycles of elaboration, presentation, questioning, and reframing to construct *frames* from *data* (Klein, Moon, & Hoffman, 2006). The Notional Model of Sensemaking in intelligence analysis has a richer depiction of intermediate products. It has a looping structure from both between and within foraging and sensemaking loops, which progressively transforms raw information to reportable results, with intermediate products of "shoe-box", evidence file, schemas, and hypotheses. These products are gradually more formal, and processes that produce each product are increasingly effortful (Pirolli & Card, 2005). The overarching idea of sensemaking is a process of iterating between less processed data and more synthesized schemas. The synthesized schemas are then used to make some downstream tasks easier, such as decision-making.

One key point of uncertainty is precisely how the models of sensemaking might inform our understanding of the complex activity of scholarly synthesis. In particular, settings in which sensemaking models have been applied or developed vary considerably, with many focused on tasks that are quite different in complexity and focus from scholarly synthesis. Examples of these diverse settings include generating retirement investment plans, creating marketing plans, constructing a new story about a topic of choice (Zhang & Soergel, 2016), developing library strategies to attract readers from certain ethnographic groups (Dervin, 1992), learning how to develop a software system in Java language (Baldonado & Winograd, 1997), or identifying missing information in a journalistic task (Attfield, 2005). These diverse end products and raw sources tend to be substantially less complex than synthesis products and scholarly sources. When sensemaking has been studied in scholarly setting, the tasks tend to be general, time-unbounded, or heavily focused foraging scholarly sources rather than synthesizing. For example, Pontis et al applied the Data/Frame Model of Sensemaking to the task of understanding the organization of a specific scientific community (Pontis & Blandford, 2015, 2016; Pontis, Blandford, Greifeneder, Attalla, & Neal, 2017). Similarly, Faisal et al. proposed a model for sensemaking a literature domain, with a vague and long-term goal of "learning and gaining knowledge of the domain" (Faisal, Cairns, & Blandford, 2006). This variation in tasks and differences from synthesis is essential to address because the way that sensemaking unfolds is task-specific: the nature of the intermediate products and the dynamics of operations on those products varies with what kind of task the sensemaking is for (Zhang & Dagobert, 2014). Depending on the task, a simple spatial clustering might suffice, or more complex and structured representations like arguments, networks, or hierarchical structures might be necessary (Faisal, Attfield, & Blandford, 2009). The complexity of the sources themselves might yield additional work or requirements not seen in the less complex topics or more structure input data from other domains.

In summary, more theoretical and empirical work is necessary to integrate advances from the domains of scholarly information practices, active reading, and sensemaking, and bridge them with the understanding we

need to support synthesis work. Because forerunning studies of scholarship and models of primitives and workflows do not highlight synthesis as a distinct process, it is unclear how that process manifests in or relates to known disciplinary, interdisciplinary practices, and the research infrastructures. Active reading research frequently focuses on the processing of individual scholarly sources and does not address the end synthesis product we care about. Finally, sensemaking theories are promising starting points for understanding synthesis; yet, most are domain-general or focused on more straightforward every-day sensemaking tasks. The output of those tasks is considerably less complex than scholarly synthesis products. Sensemaking theory itself posits that the kinds of representations that sensemakers produce are task-dependent; thus, there is uncertainty about how these models apply to the specific context of synthesis.

In this paper, we take a step towards advancing these concepts with an empirical examination of the work practices behind scholarly synthesis. Through a set of contextual interviews with scholars, we investigate the following research questions:

1. What *intermediate products* do scholars create in the course of synthesis, and for what purposes?
2. What *processes* operate on and between these intermediate products, and to what ends?
3. What are *points of friction or challenges* do scholars confront in creating and using intermediate products?

2 | METHODS

To investigate these questions, we conducted in-depth, semi-structured, contextual interviews with early-career scholars regarding their synthesis and then performed a qualitative analysis to identify emerging themes.

2.1 | Setting

Our interviews involved 10 first-year information studies Ph.D. students at a public research institution. Drawing inspiration from prior case studies (Denholm & Philpott, 2009; Holbrook et al., 2004), this sample enabled us to draw an initial understanding of the questions. While our participant sample is small in absolute numbers, there was considerable diversity amongst the participants, both demographically and in terms of their disciplinary backgrounds and interests (see Table 1). Participants' ages ranged from 23 to 37 ($M = 28$; $SD = 4.61$). Six are female, while four are male. Half of the participants finished their undergraduate education within 5 years of the study, while two finished it more than 10 years ago. Most entered the doctoral program directly from a previous academic degree program, except one (P7) who had intermediate full-time work experience. Finally, there was considerable topical and disciplinary diversity in participants' research interests. We anticipated that this diversity would allow us to both observe rich variation and idiosyncrasies in synthesis practices regarding the choices of tool adoption and knowledge management skills. At the same time, the shared context of being part of the same cohort of first-year PhD students in the same department and institution provides the opportunity for overlap and repeating patterns amongst the participants. We note also that our sample was a near-exhaustive sample from this cohort (the total number of students in the cohort was 12; two students dropped out of the study due to time considerations).

2.2 | Interview protocol

Our interviews used the more familiar language of "literature reviews" with participants to describe the task of synthesis. The interview began with a general question about the literature reviewing workflow: *"You know your regular workflow better than we do. How do you generally manage knowledge (e.g., reading papers, taking classes, etc.)?"* Then, the interview prompted participants to recall their experience of a recent synthesis effort, its development stage, and the ultimate purpose. Within that context, they described (and show samples, where appropriate) their workflow as applied to that recent literature review. In addition to follow-up questions, the interview protocol also included a series of probing questions to be asked where appropriate, such as *"Did you use any literature review tool? Why that tool?" "How do you organize your information?" "How satisfied are you with your setup/workflow?" or "Where did you encounter challenge/friction?"* Each interview lasted for approximately 1 hr, with a compensation of $20 in cash (as approved by the University's human-subjects review board). Interviews were audio-recorded, and screenshots/pictures were taken of participants' workflow components where appropriate.

2.3 | Analysis approach

We followed an iterative content-coding approach to analyzing the data from the interviews. The first and last authors iteratively went through the interview

TABLE 1 Participant biographies, research interests, and their recent syntheses

ID	Biographical info (Age, gender, and education trajectory)	Research interest	Characteristics of a recent synthesis
P1	27 years old; female; started PhD immediately after master's; transfer student after 1st year in another PhD program	Designing digital tools to help school children learn math concepts in everyday objects, drawing from the topic of embodied cognition	Brainstorming a research project with the advisor
P2	23 years old; male; started PhD immediately after undergraduate	Analyzing communication problems in online communities with a business goal	
P3	24 years old; female; started PhD immediately after master's	Studying and promoting peoples' social and emotional behaviors in VR interactions while considering their personalities and cultural backgrounds	
P4	23 years old; female; started PhD immediately after undergraduate	Augmenting and diversifying design and creative processes through the creation of novel configurations, methods, and creativity support tools	
P5	25 years old; male; started PhD immediately after master's	Designing makerspaces with texture encoding/decoding and sensory technology	
P8	28 years old; male; started PhD immediately after master's)	Understanding data tracking in video streaming services, especially on how the platform conceptualizes user identity with the data	
P10	31 years old; female; transfer student, finished 3 years in the previous institution	Educating first-generation immigrant children about their culture, specifically for Latino community in the US using gamification theories	
P6	33 years old; female; started PhD immediately after master's; transfer student after 2nd year in another PhD program	Developing a game-based curriculum for mathematics education	Preparing for a proposal
P7	37 years old; female; finished the previous master's in 2007); currently, a part-time PhD student, while working for a trusted agent and doing e-retail part-time	Understanding the role and requirements (esp. ethical) of cybersecurity for financial information and intellectual property	Writing a synthesis for a research-related course projectsover 40+ relevant literature
P9	28 years old; male; started PhD immediately after master's	Combating misinformation on social media with machine learning techniques	Preparing a paper submission

audio recordings, interviewer notes, and pictures/screenshots. The three research questions guided us to note emerging themes and observation memos from each data analysis session. The goal was to carefully describe the themes relevant to those questions as they emerged from, and were grounded in, the participants' quotes. At appropriate points, the two authors met to compare notes and memos, crystallizing their shared understanding and divergent observations in a collaborative memo.

3 | FINDINGS

This study identified different intermediate products that scholars produce as they work toward synthesis. These serve as benchmarks for scholars' progression toward synthesis and help to frame our understanding of the complex and non-linear processes they undertake to progress toward a final product. For each scholar, these processes and products are embedded in a complex ecology of tools. Analysis of scholars' idiosyncratic use of these

tools points to challenges and points of friction in the process of synthesis, some of which may be necessary and generative rather than merely disruptive.

3.1 | Three kinds of intermediate synthesis products

Our coding identified three distinctive intermediate products of synthesis work:

1. **In-source annotations**, or marginalia within sources (most often papers that align with scholars' interests);
2. **Per-source summaries**, which distilled the conceptual building blocks derived from in-source annotations;
3. **Cross-source syntheses**, which represent scholars' overall understanding of a research area based on the set of papers being synthesized.

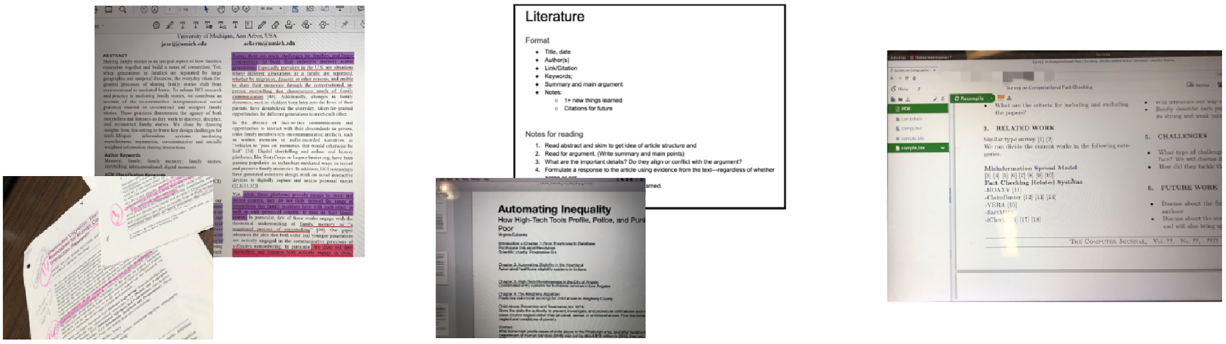
Figure 1 shows examples of each of these intermediate products. Each scholar created their products differently, relying on a diversity of organizational schemes and tools. While participants themselves did not articulate a three-part synthesis process, each participant produced a series of outputs that are well represented by these three categories. Each of their synthesis processes demonstrated a gradual progression from each product to the next (as described in section 3.2). While the three categories of intermediate products described here may be intuitive and unsurprising, they provide a useful framework for identifying the more evasive and complex processes that scholars undertake to progress toward synthesis.

In-source annotations include markings and marginal notes within a single source. Their primary function appears to be to mark elements and capture observations that align with a scholar's interest. Annotations' granularity and functional goals varied considerably, including overview arguments, key definitions, inter-connections, and specific examples. P1 simply referred to them as "fascinations."

Half of the participants in our sample mentioned highlighting as one manifestation of in-source annotations on either printed or digital documents. For example, P1 prefers to print papers out and scribbles around on the margins. As shown in Figure 1, P6 brought a set of printed papers as examples during their interviews. They read each of them for multiple iterations to decide which parts in the paper are important and need to highlight. Among participants who highlight, three use different visual styles, such as color highlights and underlines, to differentiate distinctive notions of interest. Figure 1 includes the color schemes from P10, where purple highlights are important themes, and underlines are supporting evidence for concepts. Similarly, P2 suggested a need to use different colors for the paper's central claims so that they could recognize them when revisiting the paper after several weeks.

The three who highlighted also mentioned notes or scribbles on sources as another manifestation. P10 used the comment function (the "little yellow bubbles") in their PDF viewer to create these. They explained the advantage of appending notes to papers as helping to strengthen their recall of the ideas. While the effort to scribble these simple-form notes, usually single words, was negligible, the benefits for memory for future reuse were substantial. The participant recalled an experience of searching scribbled notes more often than searching highlights, suggesting that scribbled notes were cues for memory.

Per-source summaries are written distillations of the key conceptual building blocks of a single source that the scholar might want to use as part of a synthesis. The



(a) (b) (c)

FIGURE 1 Screenshot examples of the three kinds of intermediate synthesis products: (a) In-source annotations; (b) Per-source summaries; (c) Cross-source syntheses

types included findings, theories, concepts, and solutions, as well as questions and issues sparked by the paper. For example, P9 captures what a paper does, and how they do it, in question forms such as "how they build knowledge graph from Wikipedia?" or "What is the knowledge graph representing?"

To most participants, per-source summaries manifested as a section or a paragraph dedicated to a paper. For example, P4 mentioned that they wrote a note for each paper aside from their highlights in the paper. Compared with the highlights which tend to be overused and overwhelming, the notes are much more succinct. Sometimes, the paragraphs reside as a note affiliated with an entry in reference managers like Mendeley (Zaugg, West, Tateishi, & Randall, 2011) or Zotero. P10, a participant who aspired to adopt Zotero to manage their references, has been trying their "ideal workflow of successfully foraging a paper, uploading it to Zotero, and taking notes from there. Figure 1 includes an example of the template that one participant undertook to write their per-source summaries.

Participants sometimes implement the concept type as a selection of tags that scholars attached to the paper from a more extensive tag set pertinent to their research interest. P8 has been using a tagging system over time to accumulate papers that he could use when writing on a specific topic. Participants also noted that per-source summaries often happen after in-source annotations. For example, a few participants mentioned doing multiple passes of reading a source until it emerged into the form of a per-source summary.

One participant, P9, mentioned snapshots as another type of per-source summaries. Using the image capture tool from a specific PDF viewer, they took snapshots from a source and saved it to a separate location. P9 researches in computational methods and has a recurring need to capture formulas. Capturing the picture of one specific formula from one paper was useful to them to assemble into a presentation slide deck.

Finally, **cross-source syntheses** are representations of an overall understanding of the research problem emerging from the synthesized sources. They manifested in diverse ways in our participants, such as narrative/outline summaries of key ideas across a set of sources, an annotated bibliography or shared library, or even a mind map. Figure 1 shows a working draft of an annotated bibliography from P9 shared via Overleaf, the online Latex collaborative editing platform. Participants noted that cross-source syntheses often occur as the final step in the synthesis process.

Notably, cross-source syntheses are distinct from per-source summaries and in-source annotations because they are extrinsically motivated. Scholars tend to create them not for their own internal or cognitive synthesis, learning, or sensemaking, but to communicate with collaborators or readers. Consistent with prior work on scholarly primitives (Palmer et al., 2009), annotations and summaries tended to be organic byproducts of scholars' reading and writing processes without external pressure or direction, even when scholars were not reading toward a specific synthesis goal. In contrast, cross-source syntheses only sometimes manifested as durable external representations. For example, P3 described doing a synthesis mostly for growing their own knowledge, as an exploration stage to get a sense of a domain. The synthesis they wrote was not fully externalized; P3 preferred to orally communicate their synthesis with their advisor during a weekly meeting, while the cross-source synthesis product they constructed in OneNote will rarely reach their advisor. P3 noted that their reluctance to share the cross-source synthesis was in part due to the challenge for another person to understand at a glance ("I used this to present to my advisor for one time, but it was hard for them to read in real-time. They prefer to listen to me.") In contrast, P2's synthesis process began with their advisor providing a list of paper titles as the seeds to synthesize from, to get into a funded project. (P2 later added other interesting readings to the set.). Thus, external motivations or scaffolding may influence the degree and form of the external manifestation of cross-source syntheses.

3.2 | Processes operating on and between intermediate synthesis products

Three key themes emerged from the data regarding processes operating on and between intermediate synthesis products.

First, participants described an **upward progression** of the products in increasing levels of structure and formality, transitioning from an earlier, casual, low-level phase, to a later, formal, high-level phase. Upward progression happens from in-source annotations to per-source summaries, and occasionally from per-source summaries to an overall synthesis. The information flows from a lower-level to a higher-level, usually from a more scattered form to a more integrated form. While there was a distinctive sense of progression among the products described in section 3.1., it is clear that the process is rarely linear, and that it manifests differently for each scholar as they move between products and tools.

Exemplifying the transition from in-source annotations to per-source summaries, P2 described a workflow using the Mac OS X "preview" software to read PDFs, annotating inside the paper using yellow highlights,

mainly "for personal use." Later, when they needed to export the yellow highlights into a Google Doc, to share with project collaborators, they chose not to directly copy-and-paste the yellow highlights but rather to paraphrase, as the former way "is not well-structured". P4 used the same mechanism of yellow highlighting for in-source annotation. While yellow highlights are comprehensive and overused, they kept a note as the per-source summary, which is more succinct and structured. P4 described their notes as "always (written) in the same order, sometimes with page numbers for back-reference," and would better be recorded at a consistent location, such as the "Notes" section associated with the Zotero entry.

Illustrating the transition from per-source summaries to cross-source syntheses, P3 considered the per-source summaries mainly for own use. When they needed to present with their advisor about the research topic, they would "make a more structured version in ppt, or other documents," as the more readable, externalized form of cross-source synthesis. P9 also mentioned a need to categorize altogether different papers (see Figure 1, the rightmost snapshot), as their research projects consisted of multiple types of sources: "this work is misinformation, these are checking types, these are computational approaches."

Second, participants described the deliberate construction of the intermediate products with "higher level" products in mind. They **selectively added, retained, or dropped key details** from papers that might prove useful downstream for the construction of the next intermediate product.[2] A scenario from P9 illustrates how information is selectively modified as he moves between the intermediate products. During the transition from in-source annotations to per-source summaries, P9 reads a source with two questions in mind: what this source does (the "problem") and how they do it (the "solution"). He tries to answer the two questions based on his understanding about the source. For example, while showing the interviewer a specific source he had read, he noted "this is a data mining paper that is building a knowledge graph from Wikipedia, and they are doing it by first defining the ontology of what the knowledge graph is representing." In the form of a schematized problem-solution statement, these answers become inputs to construct a structured per-source summary. These summaries then help construct the overall synthesis. P9 noted that writing a summary in this way keeps important things about a source at hand so that when he is deciding whether a source should go into the related work section in his manuscript, he could quickly recap what the source is about from the summary. In contrast, sometimes information is added to per-source summaries to facilitate cross-source synthesis. For example, P2 showed the interviewer a cross-source synthesis, manifested in a Google doc, that they were working on for a group project on bike-sharing. Collaborators shared access to the Google doc such that everyone could look at it during an online brainstorming session. While P2 was the primary person who frequently edits the document offline, doing work such as adding background papers, re-structuring sections, they made sure to include necessary context whenever a new source got referenced with a per-source summary. Instead of directly copy-and-pasting contents from the paper to form the summary, he paraphrased. In this way, collaborators could better understand the context of the cited source without stepping in earlier processes and repeating what he did. A few of our participants also mentioned being selective about what contents to highlight based on the anticipated value of the contents for reuse in subsequent sub-processes.

Finally, participants described the progression from annotations to cross-source synthesis as non-linear in that their intermediate per-source summaries were not always sufficient, requiring them to refer back to earlier in-source annotations. These **backward engagements** from the cross-source synthesis product down to in-source annotations were most common when there was a substantial time lapse between the stages. Recapitulation was achieved by connecting to the lowest-level in-source annotation or grounding the per-source summary in its context from the in-source annotation (denoted by the red arrows in Figure 2). P4 anticipated this need to back-reference and adopted a strategy of organizing the content in their per-source summary in the same order as the page numbers. Sometimes even returning to the in-source annotations and per-source summaries was insufficient because the information that was needed was not captured even at those lower levels. For example, P3 described a limitation in their current synthesis workflow, where they forget about the intentions behind writing specific in-source annotations or per-source summaries after several weeks. P3 also hypothesized that it is due to a lack of time to make important points to integrate into the notes during the initial progressing pass. P9 noted that often multiple passes through a paper was necessary, since his overall schema for the cross-source synthesis evolves, changing what details are relevant to highlight in lower-level intermediate products.

Figure 2 integrates these observations about processes with the previous descriptions of the intermediate products: the emerging model is similar in the overall structure to the Notional Model of Sensemaking (Pirolli & Card, 2005), with iterative processes producing progressively more structured and formalized intermediate products, but with frequent back-referencing.

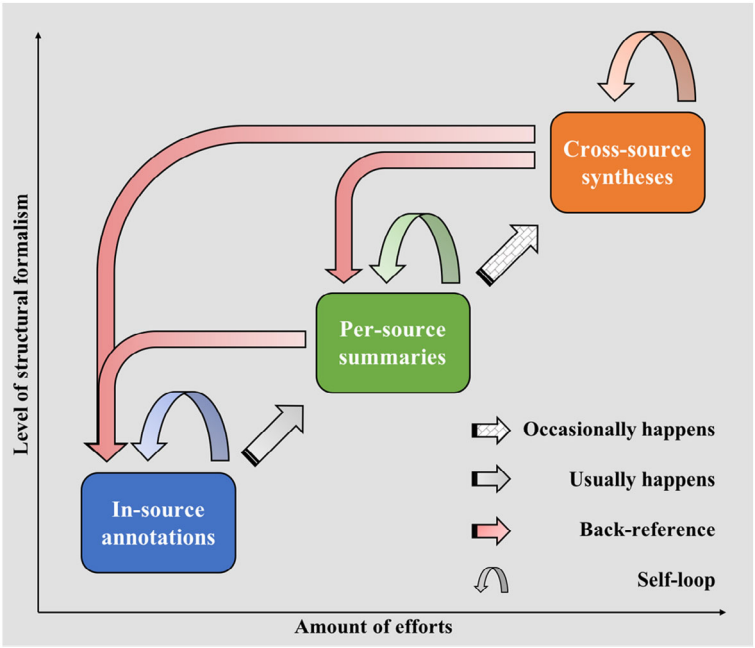


FIGURE 2 A model for scholarly synthesis that resembles the Notional Model of Sensemaking. The x-axis denotes the amount of effort to produce intermediate products. The y-axis denotes the level of structural formalization in the intermediate products. These both increase from in-source annotations to per-source summaries, and finally cross-source syntheses as denoted by the grey arrows. Within each intermediate product there are self-loops. Red arrows are backward engagements

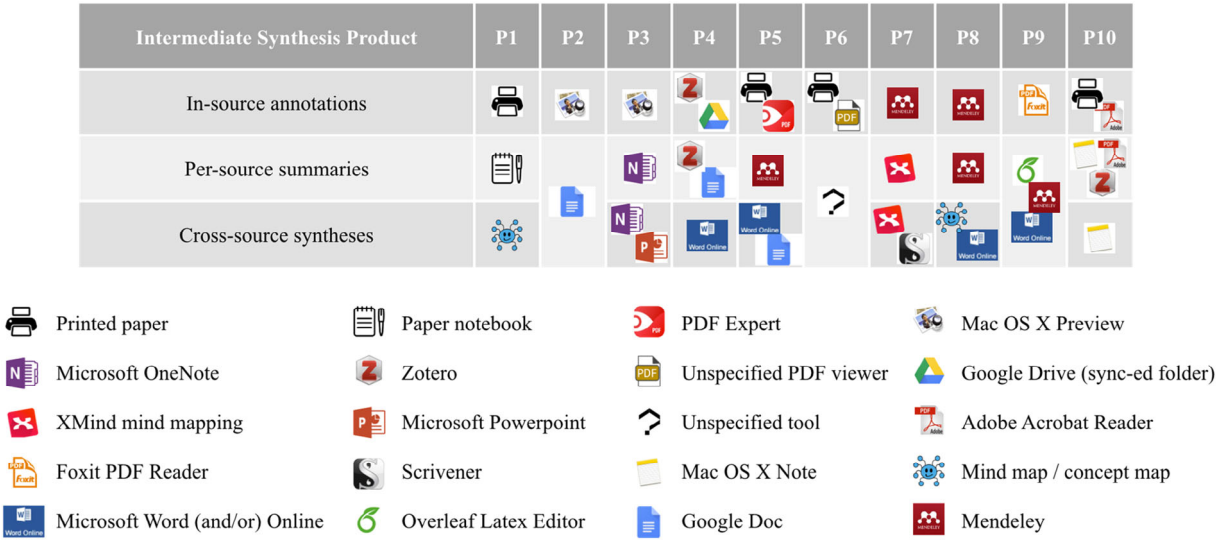


FIGURE 3 Tools used by each participant

3.3 | Intermediate synthesis products embedded in complex ecologies of tools

Having identified themes about the nature of the intermediate products, as well as the processes operating on and between the intermediate products, we now turn our attention back to the particulars of the tools in which these products and processes operate.

There was considerable diversity in the tools selected, both within the processes of individual scholars and across the whole group. As Figure 3 shows, our participants used a wide selection of tools. None of the participants used one tool to do everything; instead, each participant used 4.2 different tools on average. Likewise, not all participants followed the pattern of using one tool for one intermediate product. P2 and P9 used the same document editing software – Google Docs or Overleaf, respectively, for their per-source summary and cross-source synthesis. The document first consists of a less structured list of per-source summary, then gets organized to become a cross-source synthesis. In addition to newer technologies like document editors, PDF readers, and reference managers, many of our participants still used printed paper in conjunction with other tools, often

alongside PDF readers (P5, P6, and P10), citing benefits like enhanced concentration, tangibility, and ease of organization.

Tools were frequently appropriated (Carroll, Howard, Peck, & Murphy, 2002) to meet requirements for which they were not purpose-built to create and manipulate the intermediate products. For example, reference managers were initially used as a tool for tracking reading lists (similar to its original design purpose). But half of the participants used the note section from reference managers to write per-source summaries. P4 considers the reference manager an excellent tool to uniquely associate notes to the sources, especially for sources that are not locally saved ("Some of them don't have a PDF, but only links to online resources, but I still write notes under the Zotero file").

A key *point of friction* that scholars confront in creating and using intermediate products is tool separation. Despite some overlap in tools for intermediate products, patterns of usage in the tools also reinforced the three intermediate products' separateness. We observed this chiefly in the additional appropriation of tools' affordances to deal with decoupling between tools. The necessity of moving between numerous dedicated tools and platforms to produce different intermediate products during the overall process of synthesis was perceived as disrupting or hindering the synthesis process. Describing the friction of memorizing where and what ephemeral knowledge is, one participant (P7) noted: "so many pockets to put things in, I tend to forget." P8 mentioned several times the word "consolidate": that they wanted to "consolidate (how they do) the annotated bibliography. If it is a collaborative one, use Google drive; if it is just me, I make a folder on my local machine and use a Word document to save those ... take all those papers into a central location, easy enough to search on a local machine."

In response to these challenges, a few participants developed coping mechanisms. For example, P4 showed their desktop setup with a left-to-right simultaneous display of the first two intermediate synthesis products, with reading window for in-source annotations on the left and note-taking window for per-source summaries on the right. P5 abandoned the generic naming of the PDF in the file system, which was automatically created during the creation of a Mendeley entry. They consider it less organized "when the number of papers grows, and you have so many tabs open in your PDF viewer. You do not know how to go back (to a previous paper) when the PDF is named generically." Instead, P5 chose to rename the PDFs with keywords that they could remember, to quickly access their in-source annotations through search in the file system.

The overall picture seems to be a creatively appropriated, fluid ecology of tools, though not without significant frictions.

4 | DISCUSSION

The above findings should be seen as the first step towards a richer understanding of the distinctive and problematic scholarly product of synthesis. In particular, further work should examine how our findings generalize to scholars who are not also beginning researchers: expert scholars may have practices that vary considerably from the patterns identified here, with different implications for tool design. That said, the resemblance of our identified patterns, such as non-linear sequences of intermediate products to prior models of sensemaking, gives us hope that our findings can inform other populations'. Additionally, our methodology limits our ability to observe the practices at the moment, separate from the recall of participants. While our anchoring of interviews in a specific recent synthesis experience helps mitigate concerns about recall bias, future work could significantly enrich our understanding with observational methods like protocol analysis or extended ethnographic observation.

Despite these limitations, we believe our findings yield implications for understanding synthesis and scholarly work more generally. First, there is a question of how to interpret the findings of friction and challenges that we observed. Backward engagement with "earlier" less refined intermediate products was described in partially negative terms as a friction. The first instinct from a tool designer would be to design mechanisms to remove that friction, perhaps by capturing more information that might be selectively dropped in earlier stages but become necessary later. However, this may not achieve the goal of augmenting synthesis. Practices such as selecting things for retention, adding contextual details, revisiting prior ideas, are reminiscent of fundamental activities in scholarship and sensemaking, such as rereading (Palmer et al., 2009), or comparison and conceptual combination (Zhang & Dagobert, 2014). Thus, the described points of friction among tools and products might be necessary and productive for the process of synthesis in some cases. Creating a single end-to-end tool to support this work, or to facilitate a "seamless" synthesis experience, might remove opportunities for users to encounter obstacles that push them to do the intellectual work of synthesis. For example, the processes of scholars articulating summaries in a separate document and rationales for color-coded, in-source highlights and copying quotations may afford opportunities for them to generate new, synthetic understanding. More work is needed to carefully

distinguish frictions which are extraneous and should be removed, streamlined, or automated, and frictions which are "desirable difficulties" (Bjork, 1994) that are constitutive of synthesis work.

Second, the tools and practices identified suggest that synthesis is a foundational scholarly process that both builds upon and pervades other recognized activities across disciplines, including reading, writing, and collaborating (Palmer et al., 2009). As described above, research in information seeking and practices has long sought to develop abstract models of researchers' information activities and *primitives*, intending to guide the development of broadly useful information systems and research tools. However, *synthesis* does not appear as a distinct activity or primitive in forerunning models of scholarly practices. It may be because these models have primarily focused on empirically observable, discrete activities. In contrast, synthesis may be better understood as a pervasive, complex, socio-cognitive process (even when it generates tangible, even familiar products, like literature reviews). In particular, synthesis seems closely related to what Palmer et al. (2009) describe as "cross-cutting primitives," or activities that cut across all facets of research. Examples of cross-cutting primitives include *monitoring* ongoing and relevant developments; *notetaking* as a formative component of other processes ranging from reading to data collection, and *translating* for collaborators and audiences across disciplinary boundaries. Unlike these cross-cutting primitives, however, synthesis is generally directed toward a final goal or endpoint. It happens in a trajectory that seems to move through multiple stages before reaching that goal. By better understanding how synthesis relates to existing models of scholarly practice, researchers could align the intermediate products of synthesis work with emergent, formal models (e.g., ontologies) of knowledge representation and scholarly communication that have been developed to support next-generation research infrastructures.

Finally, our results suggest that more advances in understanding and supporting synthesis can come from integrating ideas from the relatively disparate literature on scholarly primitives, active reading, and sensemaking. In isolation, the intermediate products and practices we identified are familiar to some of these literature; however, understanding how these intermediate products and practices combine to yield barriers to synthesis will likely be a productive avenue for future work.

5 | CONCLUSION

In this paper, we sought to unpack the complex practices behind scholarly synthesis. Informed by prior theories of scholarly primitives, active reading, and sensemaking, we investigated the intermediate products, practices, and tools that scholars weave together to generate new intellectual wholes from diverse scholarly sources. Through in-depth interviews with our sample of scholars engaging in creating synthesis products, we found that our participants created and manipulated three distinctive classes of intermediate products: in-source annotations, per-source summaries, and cross-source syntheses. They employed complex and effortful practices, which generally progressed the products upward in terms of formality and structure, but in a complex, non-linear fashion. Finally, intermediate products and practices were embedded in a complex ecology of tools that are often appropriated for synthesis. Future work should examine how findings generalize to expert scholars, and observe synthesis practices with "in-the-moment" observational methods. We believe our findings yield implications for understanding synthesis and scholarly work more generally.

ACKNOWLEDGMENTS

The authors would also like to thank study participants for their interview contributions.

ENDNOTES

[1] Press release: The Prize in Economic Sciences 2019. https://www.nobelprize.org/prizes/economic-sciences/2019/press-release/

[2] A similar idea is "look-ahead modeling" in agile software development, also "backlog grooming" or "backlog refinement, which is the ongoing process of reviewing product backlog items and checking that they are appropriately prepared and ordered in a way that makes them clear and executable for teams once they enter sprints via the sprint planning activity.

REFERENCES

Attfield, S. (2005). Information seeking, gathering and review: Journalism as a case study for the design of search and authoring systems (Doctoral thesis) University of London.

Baldonado, M. Q. W., & Winograd, T. (1997). SenseMaker: An information-exploration interface supporting the contextual evolution of a user's interests. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, pp. 11–18*, Atlanta, GA. https://doi.org/10.1145/258549.258563

Bem, D. J. (1995). Writing a review article for psychological bulletin. *Psychological Bulletin*, *118*(2), 172–177. https://doi.org/10.1037/0033-2909.118.2.172

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (Vol. 185). Cambridge, MA: MIT Press.

Blanke, T., & Hedges, M. (2013). Scholarly primitives: Building institutional infrastructure for humanities e-Science. *Future Generation Computer Systems*, *29*(2), 654–661.

Carroll, J., Howard, S., Peck, J., & Murphy, J. (2002). A field study of perceptions and use of mobile telephones by 16 to 22 year

olds. *Journal of Information Technology Theory and Application*, *4*(2), 6.

Cetina, K. K. (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge, MA: Harvard University Press.

Denholm, C. J., & Philpott, C. (2009). Making the implicit explicit: Creating performance expectations for the dissertation. *Quality Assurance in Education*, *17*(2), 204–206. https://doi.org/10.1108/09684880910951408

Dervin, B. (1992). From the mind's eye of the user: The sense-making qualitative-quantitative methodology. Sense-Making Methodology Reader. Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/2281/Dervin1992a.htm

Ellis, D. (1993). Modeling the information-seeking patterns of academic researchers: A grounded theory approach. *The Library Quarterly*, *63*(4), 469–486. https://doi.org/10.1086/602622

Faisal, S., Attfield, S., & Blandford, A. (2009). A classification of sensemaking representations. *CHI 2009 Workshop on Sensemaking*, Boston, MA.

Faisal, S., Cairns, P., & Blandford, A. (2006). Developing user requirements for visualizations of literature knowledge domains. *Proc. IV07, pp. 264–269*, London, England.

Froese, A. D., Gantz, B. S., & Henry, A. L. (1998). Teaching students to write literature reviews: A meta-analytic model. *Teaching of Psychology*, *25*(2), 102–105.

Gradmann, S., Hennicke, S., Tschumpel, G., Dill, K., Thoden, K., Pichler, A., & Morbidoni, C. (2015). *Beyond Infrastructure! Modelling the Scholarly Domain*. EU funded DM2E (pp. 1–44).

Granello, D. H. (2001). Promoting cognitive complexity in graduate written work: Using Bloom's taxonomy as a pedagogical tool to improve literature reviews. *Counselor Education and Supervision*, *40*(4), 292–307.

Holbrook, A., Bourke, S., Lovat, T., & Dally, K. (2004). Investigating PhD thesis examination reports. *Doctoral Assessment in Australia*, *41*(2), 98–120. https://doi.org/10.1016/j.ijer.2005.04.008

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems*, *21*(4), 70–73.

Morris, M. R., Brush, A. B., & Meyers, B. R. (2007). Reading revisited: Evaluating the usability of digital display surfaces for active reading tasks, *Proceedings of IEEE Tabletop 2007*, *IEEE, pp. 79–86*, Newport, Rhode Island.

O'Hara, K. (1996). Towards a typology of reading goals.

O'hara, K., & Sellen, A. (1997). A comparison of reading paper and on-line documents, *Proceedings of CHI, pp. 335–342*, Atlanta, GA.

Orr, J. E. (1996). *Talking about machines: An ethnography of a modern job* (1st ed.). Ithaca, NY: ILR Press.

Palmer, C. L., Teffeau, L. C., & Pirmann, C. M. (2009). Scholarly Information Practices in the Online Environment: Themes from the Literature and Implications for Library Service Development (Report Commisioned by OCLC Research, p. 59). Retrieved from http://www.oclc.org/programs/publications/reports/2009-02.pdf

Pirolli, P., & Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *Proceedings of International Conference on Intelligence Analysis, Vol. 5, pp. 2–4*, McLean, VA.

Pontis, S., & Blandford, A. (2015). Understanding "influence:" an exploratory study of academics' processes of knowledge construction through iterative and interactive information seeking. *Journal of the Association for Information Science and Technology*, *66*(8), 1576–1593.

Pontis, S., & Blandford, A. (2016). Understanding "influence": An empirical test of the D ata-F rame T heory of S ensemaking. *Journal of the Association for Information Science and Technology*, *67*(4), 841–858.

Pontis, S., Blandford, A., Greifeneder, E., Attalla, H., & Neal, D. (2017). Keeping up to date: An academic researcher's information journey. *Journal of the Association for Information Science and Technology*, *68*(1), 22–35. https://doi.org/10.1002/asi.23623

Puckett, J. (2011). *Zotero: A guide for librarians, researchers, and educators*. Chicago, IL: Assoc of Cllge & Rsrch Libr.

Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). The cost structure of sensemaking. *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, pp. 269–276*, Amsterdam, The Netherlands. https://doi.org/10.1145/169059.169209

Strike, K., & Posner, G. (1983). Types of synthesis and their criteria.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge, England: Cambridge University Press.

Tashman, C. S., & Edwards, W. K. (2011a). Active reading and its discontents: The situations, problems and ideas of readers. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2927–2936*, Vancouver, Canada. https://doi.org/10.1145/1978942.1979376

Tashman, C. S., & Edwards, W. K. (2011b). LiquidText: A flexible, multitouch environment to support active reading. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 3285–3294*. https://doi.org/10.1145/1978942.1979430

Unsworth, J. (2000). Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This.

Vilar, P. (2015). Information behaviour of scholars. *Libellarium: Journal for the Research of Writing, Books, and Cultural Heritage Institutions*, *7*(1), 17. https://doi.org/10.15291/libellarium.v7i1.194

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, *26*, xiii–xxiii.

Zaugg, H., West, R. E., Tateishi, I., & Randall, D. L. (2011). Mendeley: Creating communities of scholarly inquiry through research collaboration. *TechTrends*, *55*(1), 32–36.

Zhang, P., & Dagobert, S. (2014). Towards a comprehensive model of the cognitive process and mechanisms of individual sensemaking. *Journal of the Association for Information Science and Technology*, *65*(9), 1733–1756. https://doi.org/10.1002/asi.23125

Zhang, P., & Soergel, D. (2016). Process patterns and conceptual changes in knowledge representations during information seeking and sensemaking: A qualitative user study. *Journal of Information Science*, *42*(1), 59–78. https://doi.org/10.1177/0165551515615834

How to cite this article: Qian X, Fenlon K, Lutters W, Chan J. Opening up the black box of scholarly synthesis: Intermediate products, processes, and tools. *Proc Assoc Inf Sci Technol*. 2020;57:e270. https://doi.org/10.1002/pra2.270