Supporting Storytelling With Evidence in Holistic Review Processes: A Participatory Design Approach

RONALD A. METOYER, University of Notre Dame, USA
TYA S. CHUANROMANEE, University of Notre Dame, USA
GINA M. GIRGIS, Troy University, USA
QIYU ZHI, University of Notre Dame, USA
ELEANOR C. KINYON, University of Notre Dame, USA

Review processes involve complex and often subjective decision-making tasks in which individual reviewers must read and rate submissions, such as a college application, along many relevant dimensions and typically with a rubric in mind. A common part of the work is *committee review*, where individual reviewers meet to discuss the merits of a particular submission in order to recommend an accept or reject decision. Prior work indicates that visualization and sensemaking support may be beneficial in such processes where reviewers must present the "story" of the applicant under question. We conducted a series of participatory design workshops with reviewers in the domain of holistic college admissions to better understand the challenges and opportunities regarding storytelling. Based on these workshops, we contribute a characterization for how reviewers in this domain construct visual stories, we provide guidance for designing for evidence capture and storytelling, and we draw parallels and distinctions between this domain and other reviewing domains.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; **Field studies**; *Information visualization*;

Keywords: participatory design; holistic review; storytelling; sensemaking

ACM Reference Format:
Ronald A. Metoyer, Tya S. Chuanromanee, Gina M. Girgis, Qiyu Zhi, and Eleanor C. Kinyon. 2020. Supporting Storytelling With Evidence in Holistic Review Processes: A Participatory Design Approach. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 4, CSCW1, Article 61 (May 2020). ACM, New York, NY. 24 pages. https://doi.org/10.1145/3392870

1 INTRODUCTION

Critical review is an important process in many domains including holistic admissions, hiring, and grant funding, to name a few familiar examples. These domains, in turn, carry important societal, political, and economic implications.

Consider, for example, the recent spotlight on college admissions processes. A 2018 lawsuit filed on behalf of Asian-American students against Harvard University alleged that Harvard discriminated against Asian-American applicants [29]. The trial thrusted holistic review into the spotlight,

Authors' addresses: Ronald A. Metoyer, rmetoyer@nd.edu, University of Notre Dame, Notre Dame, Indiana, USA, 46556; Tya S. Chuanromanee, tchuanro@nd.edu, University of Notre Dame, Notre Dame, Indiana, USA, 46556; Gina M. Girgis, ggirgis@troy.edu, Troy University, Troy, Alabama, USA, 36082; Qiyu Zhi, qzhi@nd.edu, University of Notre Dame, Notre Dame, Indiana, USA, 46556; Eleanor C. Kinyon, ekinyon@nd.edu, University of Notre Dame, Notre Dame, Indiana, USA, 46556.

2573-0142/2020/5-ART61 $15.00
https://doi.org/10.1145/3392870

scrutinizing the specific practices of Harvard and likely general practices among universities that employ some form of holistic review. Admissions processes were again highlighted in 2019 when many wealthy individuals, totaling 50 in six states, were accused by the Justice Department of participating in a major college admissions scandal that rocked many elite universities [2]. These are just two examples in recent years that highlight the importance of review processes, particularly in higher education.

These processes are complex for many reasons. In holistic admissions review, for example, admissions officers (i.e., reviewers or readers) must read hundreds or thousands of applications under tight time constraints. For each applicant, the reviewer undertakes a complex task that mirrors many parts of sensemaking [28], where she must gather evidence in support of a student that she would like to recommend for admission. This evidence may come in many forms including quantitative test scores, non-academic accomplishments such as sports or hobbies, and subjective impressions found in letters of recommendation or personal essays [35]. This requires that the reviewer keep track of numerous facts as well as their own reasoning processes and can be a significant cognitive challenge[38]. To complicate matters, reviewers must then typically present that applicant's case during a committee meeting that may take place several days later when it is likely that many of the facts and rationale have long since faded from memory. They typically rely on a representation called a "Decision Sheet" which is essentially a summary (or rubric) of the important facets of the application. The *Decision Sheet* is generally a static representation, completely textual in nature, and is auto-populated with application data and reviewers' notes [38].

In a sense, the *Decision Sheet* is a kind of group workspace [23] in that it allows readers to see each others' brief notes and conclusions, but lacks a representation of rationale or ways to relate pieces of information to each other or to build on others' work. This *Decision Sheet* is then used as the primary point of information during committee review meetings where reviewers present an applicant's case. Prior work indicates that there are opportunities to support the review process through sensemaking support such as visual representation and storytelling support [5, 38].

Because of the wide application [12, 44] and power [33] of visual storytelling, in this work we set out to explore how visual storytelling support might be integrated into review processes to better support reviewer practices with the ultimate goal of reducing cognitive load on reviewers and creating a more transparent and rigorous system for presenting an applicant's case. In this paper, we investigate three research questions that are relevant to this overarching goal and focus on the specific domain of college holistic review:

- RQ1: How might admissions officers construct applicant stories from evidence artifacts?
- RQ2: How do evidence-driven stories map to existing genres for visual narratives?
- RQ3: How might we design for storytelling in such situations?

We contribute a nuanced understanding of the committee review process in a college admissions holistic review setting, a characterization of the kinds of stories constructed by reviewers in this domain, and design suggestions for how to support holistic review storytelling and sensemaking more generally.

2 BACKGROUND AND RELATED WORK

To get a better understanding of the potential design scope for the holistic admissions domain, we discuss background literature in admissions and holistic review. We then briefly present relevant efforts to understand and support group or collaborative decision making.

2.1 Holistic Review

One popular form of college admissions practices is the holistic review which has been studied in several domains [1, 6, 8, 10] and explained in several non-fiction works [13, 35]. Holistic review is largely subjective in nature and requires that reviewers consider the applicant in context. This means that an applicant is rated not simply on their scores or grades, but also on those aspects with respect to the context in which they were achieved. This context may be broad and include elements such as high school, family background, adversities faced, rigor of curriculum, grade trends, non-academic accomplishments and personal qualities of the applicant. Those non-academic accomplishments might include community service, music ability, special talents, and leadership skills, for example.

Talkad Sukumar et al. studied the holistic review process used in American universities through the lens of human-computer interaction and technological support [38]. They describe a system in which application reviewers use their experience as well as a set of criteria aligned with the university's mission and goals to assess an application based on a variety of factors that range from academic performance through adversities faced to personal characteristics. When an applicant is not a clear accept or clear reject, reviewers defer applications to be discussed within review committee meetings that include multiple reviewers where the lead reviewer presents the applicant and makes a case for or against their admission. Talkad Sukumar et al. found that holistic review, while subjective, complex, and nuanced in nature, is weighted down by "the numerous variables involved and the reviewers' thought processes," yet shaped by the tools reviewers use. The digital tools used, however, are primarily textual and representational, supporting capture of notes within the application and codes of various forms, but not explicitly designed to support the more nuanced sensemaking process that reviewers actually engage in. In particular, there is little support for evidence capture and recall, rationale representation, or storytelling.

Our work is based on this initial ethnographic exploration [38] and looks particularly at extending their findings regarding the "committee review meeting" aspect of holistic review. We build upon their findings, extend our understanding of committee review, and further explore how reviewers use storytelling during these meetings.

2.2 Collaborative Review and Decision Making

Collaborative review and decision-making processes have been studied for many years, resulting in numerous publications on the topic. Of particular interest to our work are explorations into evidence capture and representation to support collaboration.

The admissions holistic review process is highly collaborative and memory intensive. Rule et al. highlight the variety of cognitive and collaborative structures that, in turn, influence recall and reconstruction of mental context [31]. From their perspective, holistic review would be divided into execution (reviewing an application) and evaluation (presenting findings and getting feedback from peers). It is clear that these two phases are temporally far apart, which Rule et al. describe as a "loose execution-evaluation loop." This structure requires that application reviewers spend time scanning applications *after the initial review* to construct and recall the relationship of different parts of each application in preparation for committee review.

In support of processes such as this where the time between review and discussion can present cognitive challenges, several have proposed systems that capture and exploit histories of usage [14, 31]. Inspired by this idea, we explore visual histories not only as a mechanism for prompting memory but for making rationale concrete.

2.3 Storytelling and Decision Making

Rule et al. emphasize that technology can facilitate storytelling by using visual histories of users' work as prompts for stories [31]. Storytelling is an essential component of the social aspect of intelligence analysis as the members of one professional group share their individual perspectives and recommendations in order for the group to collectively make a decision. This description matches that presented in [38] for holistic review committee meetings. The representation of applications during those meetings often entails storytelling in some form or another. We explore some of the relevant literature in the domain.

Hossain et al. challenge the assumption that story construction and storytelling diminish the rationality of decision making in any way. They explain how in group decision-making discussions, topics will progress during the discussion to make a decision and/or come to a conclusion [16]. They claim that the topics that the group elaborate on the most through narrative were ones involving possible group decision. The decision makers read between the lines to connect the dots and form hypotheses. This description matches the analysis and information extraction processes in the admissions domain as application reviewers strive to see through the mostly textual application to determine "fit" of the applicant to the university. By extension, the application review committee's use of stories is beneficial, and therefore it is important to employ the appropriate tools to support the representation and telling of stories.

Likewise, Xiao et al. asserted that story construction aids rationalization and sensemaking as participants tend to consider different perspectives and important arguments [42]. They elaborate on the group's behavior even further by outlining a "Story Model." This claim basically states that narrative construction is an effort to integrate available evidence and arguments into a representation to facilitate decision making. We are particularly interested in this paper in representations dominated by a rubric or set of criteria over which the applicant is evaluated.

Mentis et al. provide what could be viewed as an extension to the claims made by Hossain et al. about storytelling validity and its effect on rationality in group decision making by qualifying that group decision making itself is not a rational process and therefore could use the support of an easy-to-communicate representation of collected evidence to bolster the argument being made by the group. Both Rule et al. [31] and Mentis et al. [24] suggest an annotation tool to show previously-captured rationale. We explore this approach in our participatory design workshops.

2.4 Group Decision Support Systems and Computer-Supported Argument Visualization

Group Decision Support Systems (GDSS) [17] and related Computer-Supported Argument Visualization (CSAV) tools have been studied for many years [40]. Most relevant to our work are efforts that focus on the benefits of "information pooling" for visualizing arguments for or against a particular decision option, especially in mitigating biases. Our recommendations are inspired by this work and suggest a visual solution for presenting the pros and cons of an applicant during holistic committee review.

GDSS are split into two types, those that only support the decision-making process through book-keeping and visualization or building models, and those that use mathematical algorithms to predict or suggest outcomes, which alter the deliberation of user groups. The latter is met with skepticism and Introne suggests that there is not enough evidence to prove the benefits of employing these types [17]. Instead, he emphasizes the importance of the system having an easy way to characterize "arguments" as a pro or con to an "alternative." This perspective inspired design ideas explored in our participatory workshops.

Van Bruggen et al. [39] introduce a framework for CSAV as a whole. Of interest is their definition of representational notations, representational tools, and representational artifacts. Representational notations are discrete objects that are components in arguments and are governed by rules. The software that show these notations are called representational tools. They differ in how they choose to implement and display the notations. Representational artifacts communicate how the notations relate to each other. Examples include argument maps or diagrams. The authors emphasize that CSAV tools should balance the specificity of a problem with ease of use. In this paper, we explore a design of a representational tool that shows representational notations in a simple but intuitive fashion, with support for representational artifacts.

Previous efforts to utilize visualization to support group decision making include Alonso et al.'s [4] tool that presents the group's current consensus using diagrams to aid decision making. In addition to the current consensus, they show which of the current experts in the group have consistent or different opinions. Along the theme of using visualization to show the structure of a group and one's place in that group, Jongsawat and Premchaiswadi [18] presented a web-based GDSS system that uses the Nominal Group Technique, a special structured decision-making methodology. Their GDSS system shows each group member's disposition and status to support online collaboration.

Palomares and Martínez presented tools using visualizations to support decision making in contexts with high uncertainty [26] and group decision-making in general [27]. Another similar interface is ConsensUs, a visualization tool that displays all members' opinions individually and in aggregate [22]. The benefit of ConsensUs is that it helps users identify the points of contention while also clearly exposing participants to others' points of view. However, like the application reading process, a user can see other users' current ratings when they are forming an opinion on a topic, which can lead them to alter their answers. Holistic admissions review also differs in that most applications are read by a very small number of reviewers (typically 2-3).

Walton focuses on utilizing visualization to support argumentation by evaluating three different tools using witness testimony as an application and emphasizing corroborative (or secondary) evidence [41]. Walton distinguished between two types of arguments regarding secondary evidence, supportive and convergent. In supportive arguments, or linked arguments, each piece of evidence connects to one another in reaching the conclusion. In convergent arguments, each piece of evidence stands on its own to support the conclusion, that is, the pieces of evidence are not dependent. In the case of holistic admissions, both kinds of arguments are used. For example, a reader who is telling a story connects parts of an application to each other to support a conclusion that this student should be accepted. Readers may also pick different application components that are independent of each other to argue, for example, that a student's performance is too weak and the student should not be accepted to the university.

In total, this prior work investigating holistic admissions, collaborative review, storytelling, and GDSS suggests that visual narratives and sensemaking support could be helpful during the committee review process, in particular, to alleviate memory challenges, increase efficiency, and improve group decision-making. In the following sections, we explore this possibility.

3 PARTICIPATORY DESIGN WORKSHOPS

Gaining consistent access to stakeholders is a challenge in the holistic review domain due to the frequent travel, extreme time constraints, and deadline-driven nature of the process throughout much of the academic year. The sensitive nature of the work also precludes observation of real practices in action.

Given the nature of the domain, we chose to employ a participatory design (PD) approach over multiple workshop sessions during the summer, the least busy time of the year for admissions

officers. Our goal was to iteratively refine our thinking and ideas *in collaboration* with admissions officers over these three 2-hour PD sessions. The following section describes each participatory design session. For each session, we outline our process, data collection, analysis, and preliminary results that were used to inform the subsequent sessions and ultimately inform our design recommendations. Rather than discuss the results of each session in aggregate at the end of the paper, we have chosen to discuss each session, in turn, to illustrate how it informed future sessions.

We conducted three participatory design workshops over a six-month period that was embedded within a longer-term engagement of regular interactions and engagement of over one year with our PD collaborators. We carefully considered our workshops in an attempt to create the hybrid shared space described by Muller et al. [25] between our team of computer scientists and admissions officers. Each workshop took place in a conference room in the admissions office building — a room that is often used for the committee review process. The proposed space would, presumably, help the officers recall important procedural context while also facilitating explanation of their work to our team of scientists. We augmented that space with design materials necessary for each phase of the workshops. All workshop sessions were audio-recorded and transcribed for analysis. These workshops were approved by the University of Notre Dame's Institutional Review Board.

3.1 The PD Team

Our participatory design team was composed of a faculty member, a graduate student, and an undergraduate student, all from a computing background, and seven admissions officers from a highly selective private university that utilizes a holistic admissions review process. From a brief demographic survey administered during the first workshop we collected the admissions officers' professional background (e.g. years of experience at this particular university, years of experience in admissions in general, highest degree attained, degree field). Five of seven reviewers were alumni of the university. All of the reviewers had a Bachelors degree, with three of them also having a Masters degree. The reviewers had 3 to 18 years of experience in working in admissions and nine months to eight years of experience in the current university admissions position. We compensated each participant with a $100 Amazon gift card.

3.2 Workshop 1: Challenges and Opportunities

The first workshop was designed primarily to complete the necessary administrative processes (e.g., informed consent), set up the hybrid third space of collaboration, and characterize the challenges and opportunities of the review committee meeting. We were interested in confirming previous findings from [38], understanding if there were additional challenges that represented opportunities for design to enhance the review process, and exploring the use of evidence artifacts in storytelling.

Workshop Activities. After brief introductions and completing the informed consent documents, we began audio-recording the session. We first established some common language around participatory design and introduced the team to our expertise area through a discussion of visualization and narrative visualization in particular. We then collected initial information through a brief paper-based demographic survey to obtain an overview of the admissions officers' professional background (e.g., years of experience at this particular university, years of experience in admissions in general, highest degree attained, degree field). Then, we embarked upon a group discussion around the committee meeting review process.

To end the session, we introduced an activity inspired by the use of photo-narrative and photo-documentaries for bringing design teams into unfamiliar or new spaces [25]. This activity helped bring our research team into the admissions officers' new space and resulted in artifacts for exploration of storytelling in Workshop #2. We asked our admissions officers to read an application (preferably one that they had not read before or one that they did not remember) and while reading

it, capture, with a photograph or screenshot, anything that they considered important in making their recommendation for the applicant. The photographs could be actual photographs or snapshots of digital documents, applications, physical documents, or physical objects. We provided them with instructions on capturing and uploading the photographs into a folder hosted in a cloud service.

Data Analysis and Results. We transcribed the recordings from the PD session discussion. From transcriptions and notes, we identified representative statements of the discussion and used affinity diagramming [15] to generate themes from the session. The themes we identified largely confirmed previous findings from [38] while providing additional nuance to those findings. In particular:

- *Committee meetings are a very collaborative effort.* Working as a group levels out biases and every reader has a role during the meeting. In particular, the lead reader drives the conversations while the others serve to 1) record notes and 2) find corroborating or contradicting evidence. Readers bring different perspectives and background knowledge. Although they detailed the committee meeting setup such as the presence of a driver and a single presenter telling the story of a student, we emphasize that the committee meeting process is more collaborative than previously thought based on reports from [38]. The readers mentioned several times how working with others with a different background gives them different perspectives on the same student and levels out biases they may have for a certain student. According to P3, *"It can sometimes be frustrating because we all have different approaches. However, I think the beauty of committee levels out our personal biases. We are all more familiar with our own territories. We all have our stories and backgrounds, and we are more sympathetic to those who are more like us. At the end of the day, the idea is that all of our best characteristics and ideas and biases are brought together."*
- *Evidence collection and memory are challenges but are critical.* The information that readers collect and remember affects the stories that they tell, which then affect the recommendations of the committee. The challenges that this information collection and recall pose due to time constraints are greater than initially expected, but the amount of difficulty involved varies from reader to reader.
- *Stories have somewhat of a structure.* Certain pieces of information (e.g., test scores, GPA) are necessary parts of every applicant's story, resulting in a somewhat standardized presentation of information. However, there is also a subjective component to each story where reviewers weave a compelling narrative for the application under discussion, highlighting an applicant's unique attributes in a specific context.
- *Context is critical.* Our findings reaffirm the importance of context and give more detail to Talkad Sukumar et al.'s [38] findings that reviewers take high school context into account. We also find that readers consider environmental context as well. Readers take context into consideration with respect to aspects such as school and extracurricular activities, or the neighborhood, home, and socioeconomic conditions. Individuals cannot be viewed in the same way because each applicant has been exposed to a different context that has shaped his or her life and afforded certain privileges or disadvantages.
- *Committee must be efficient.* The number of students that readers need to discuss at each meeting is significant and readers cannot take too long to present a single student. Our findings also support the fact that time constraints are a challenge for reviewers [38]. The number of students that each reader brings to the table in a committee meeting is based on the frequency of committee meetings. Regardless, these meetings must be efficient due to the sheer number of student applications that generally must be discussed.

One new theme emerged regarding the role of the committee meetings. This theme included comments about the various ways in which *teaching* happens during committee. While the committee meeting's primary objective is to come to an agreement on a recommendation, this collaborative and transparent process is a good "teaching mechanism" where new readers learn what to look for in an application (e.g., context) and how to communicate it in a compelling and convincing fashion. It also helps readers calibrate not only their ratings but also the amount of time and effort put into each review. Reviewers learn what evidence equates to particular scores and how much time their counterparts spend in evaluating an applicant. All of these training benefits ultimately lead to a more consistent and efficient reviewing process.

Discussion of Workshop #1. The resulting themes, when considered together, lead to several interesting findings around the characteristics and challenges of committee review. As a whole, these themes can be summarized into three insights:

- Preparing for committee review is challenging due to time and memory constraints.
- The review team collaborates to efficiently and fairly evaluate an applicant, creating opportunities for training and calibration along the way.
- Committee review cannot truly be separated from individual application review where evidence is initially collected and where reviewers begin preparing to tell the student's story.

The collaborative nature, combined with cognitive memory challenges present a situation where availability bias can play a significant role. Availability bias is a cognitive bias where easily-recalled information carries a higher weight than information that is more obscure to the decision-maker [30]. Likewise, the roles of the members present opportunities for confirmation bias to both present itself (e.g., in the lead reviewer's story) and to be mitigated (e.g., by the other committee members). For the second workshop, we decided to focus on mechanisms to support the reviewers' evidence-collection process and how they might present a student's story during committee with explicit representations of evidence. Previous research [38] informed our decision to adopt visual storytelling as they have found that it could be a potentially useful tool to enhance the holistic review process. We leveraged the data collected in the photo-narrative activity to explore how reviewers generated stories from this evidence (RQ1).

3.3 Workshop 2: Evidence Collection and Storytelling

From the first workshop, it was clear that the cognitive load involved in remembering important details was a challenge in preparing for committee meetings. Driven by this challenge and inspired by the potential of visual histories to support recall [32], we set out to explore the potential role of visual histories in review processes during the second workshop. In particular, we sought to understand how officers would utilize visual snapshots of the application as evidence to construct visual stories.

Workshop Activities. We began the second workshop with a photo-narrative activity. We asked our team of reviewers to construct a collage or photo-narrative of an applicant using the snapshots and photographs that they collected during the exercise given to them at the end of Workshop #1. We asked them to focus on how they would make a recommendation and sell it to the committee, since we wanted to know how the readers interpret and organize the evidence for admitting a student.

Each participant was given a large sheet of paper (poster board size) as a canvas, along with the snapshots/photographs themselves and various other creativity tools such as markers, sticky notes, and tape. The readers were instructed to organize the photographs however they wished and to feel free to add content (e.g., annotations, labels) as they deemed appropriate. After each participant finished constructing their collage, we invited them to tell the entire team about their

photo narrative. All session discussions were audio recorded and photographs were taken of the resulting photo narratives. Figure 1 shows a recreation of a participant's collage with all sensitive information removed from the boxes that correspond to specific application data. Each of the boxes represents a snapshot that the reviewer took and is a piece of evidence taken from various parts of the application. The labels and annotations are in the reviewers' own words.

Table 1. Characterization of the photographs with respect to the Segel and Heer's classifications.

Photo Narr.	Relevant Visual Narrative Genre(s)	Visual Narrative Tactics	Narrative Structure Tactics
P1	Partitioned Poster and Flow Chart.	Visual Structuring (Arrows, clusters)	Messaging (Captions/Headlines)
P2	Flow Chart	Visual Structuring (Arrows, Clusters)	N/A
P3	Partitioned Poster	Visual Structuring (Clusters)	Messaging (Headlines)
P4	Partitioned Poster	Visual Structuring (Clusters, Checklist progress), Highlighting (Feature Distinction)	Messaging (Headlines, Annotations)
P5	Partitioned Poster and Comic Strip	N/A	Messaging (Captions, Summary/Synthesis)
P6	Flow Chart	Visual Structuring (Arrows)	Messaging (Headlines, Annotations, Summary/Synthesis)
P7	Slide Show and Partitioned Poster	Visual Structuring (Clusters)	Messaging (Headlines)

Data Analysis and Results. We used an open coding approach [36] to analyze the resulting photo narratives. We coded both the photo narratives themselves and the reviewers' presentations and discussions of their narratives. To code the photo narratives for story elements, we abstracted them to include the components and layout but not the details of what was in the photographs themselves. In developing the codebook, we focused on the primary content elements of the readers' stories, the temporal and spatial organization of the narratives, and the genres of narrative visualization as characterized by Segel and Heer [33].

Once we agreed on a codebook, two members of the design team worked together to code all seven photo-narratives and their transcriptions. Because the coding was done side by side and collaboratively, we were able to reach a full agreement on which codes were applied without the need to use a measure of inter-rater reliability. We also identified Segel and Heer's three characteristics of visualization narratives: genre, tactics, and narrative structure [33] as described in Table 1 and illustrated in Figure 2. Their categorizations are meant to characterize the design space in narrative visualization [33]. Here, visual narrative genre refers to visualizations that show the flow of information. Visual narrative tactics help facilitate the narrative and include devices that perform functions such as the structure and highlighting of visual information. Narrative structure tactics perform the same function as visual narrative tactics but are composed of non-visualization methods. Not all of the visual narrative tactics and narrative structures can be applied to the photo narratives due to their physical and non-interactive nature.

From our codes, we identified several salient themes. We found that five readers presented their narratives using clusters or subgroups that largely aligned with the *Decision Sheet*. They all included the students' academic information, demographic information, and other non-academic information such as extracurricular activities. There were nine instances of the code "overarching

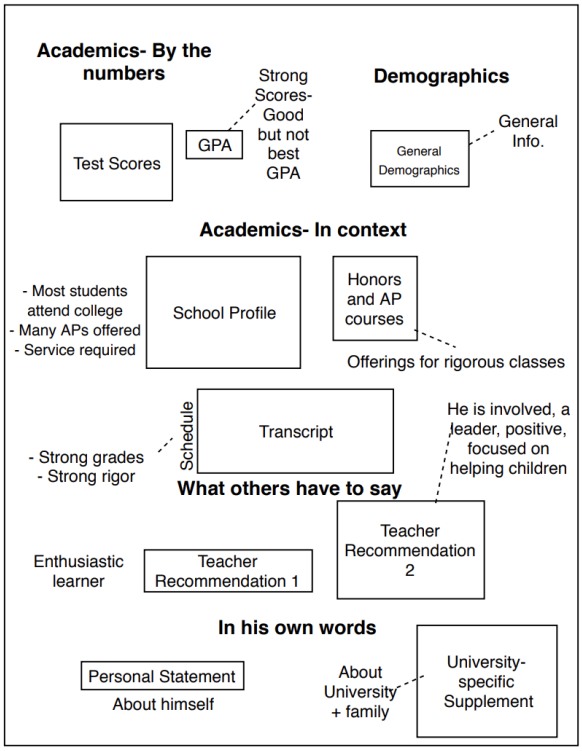

Fig. 1. P3's photo-narrative representation. This partitioned poster includes Messaging in the form of Headlines (e.g., Demographics) and Captions (e.g., Enthusiastic learner) and uses clusters as a Visual Structuring narrative tactic. See [33] for more examples of visual narrative and structure tactics.

question," which corresponds to a broad question that the reader is trying to answer by looking at the evidence they have gathered in favor or against the student.

In the collages, we found that five out of seven of the readers presented a linear ordering of information, one did not prescribe a specific ordering at all, and one presented an ordering that had multiple possible paths [33]. All included highlighting and six had messaging in the forms of labels, comments, and headlines.

From their explanations, we also identified some themes that stood out. There were 21 instances where the readers assigned some weight or importance to a particular piece of evidence or section of the photo narrative. This can be positive or negative, depending on what is discussed. For instance, P4 said, *"If the student doesn't have the academic strength and profile that we're looking for, really a lot of the other information's not going to be as relevant because it's going to stop there in terms of their candidacy."* If something stands out to the readers, it adds weight to the readers' argument. P7 said, *"We highlight very strong or very weak letters of recommendation. This is a very strong letter. Saying that if she wants her children to be like this person. And I think that's the biggest thing to say that you want your own kids to be like her."*

As also noted in [38], readers make a fair amount of inference. They connect certain pieces of information to make predictions or answer questions such as *"Is this student going to succeed here?"* [P5] or *"How does this student engage in the classroom?"* [P6]. P4 explained how he connects the dots, *"I oftentimes find myself in situations where I want to put my notes on the application specifically*

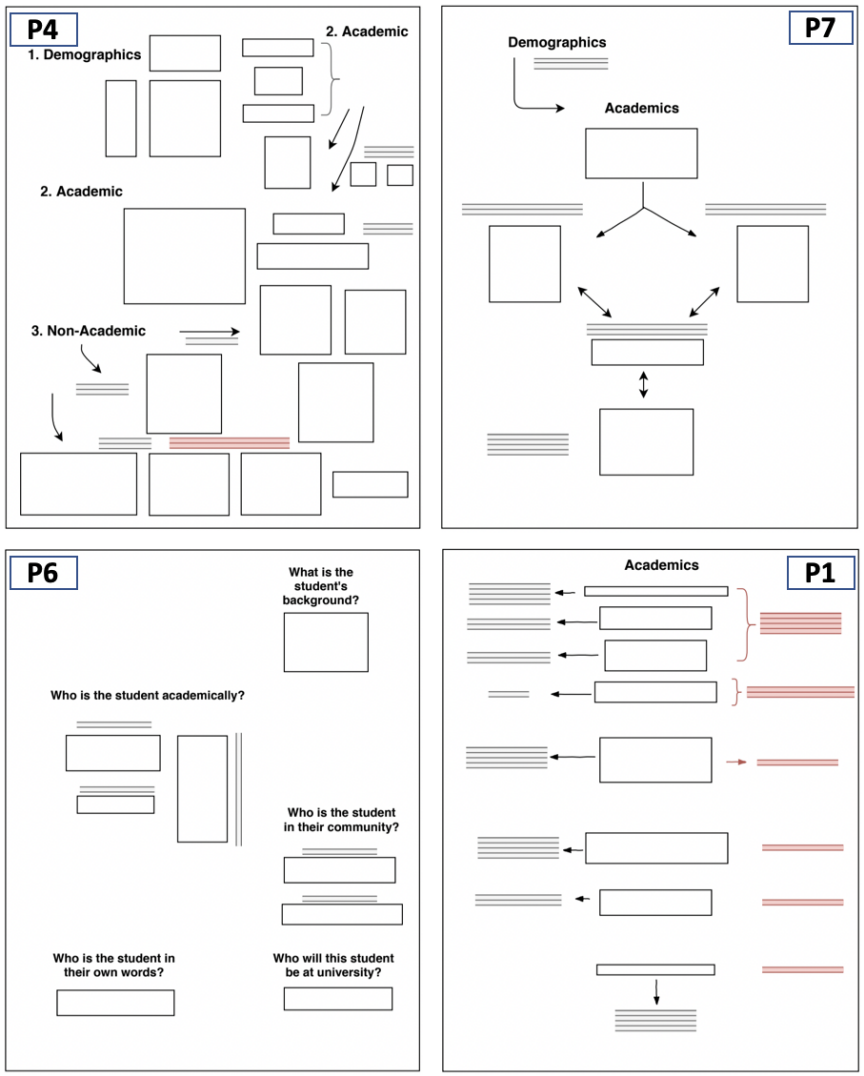

Fig. 2. Top Left: P4's photo narrative with elements of a partitioned poster as well as flow chart indicated by the arrows in the Academic and Non-Academic sections. Top Right: This photo narrative by P7 represents a clear flow-chart with multiple paths. The reviewer described it as as such, allowing for his story to unfold in multiple ways and to circle back given the bi-directional flows. Bottom Left: P6's photo-narrative is another example of a partitioned poster organized by overarching "questions" that drive the story. Finally, the photo-narrative at the bottom right by P1 contains elements of the partitioned poster as well as the comic strip given the "commentary" captions shown in red.

about something. Some system in place to [incorporate other parts of the student story]...The first thing that comes to mind is a student who started their own company. That comes immediately to me as I go on and search, can I find out what kind of company? Can I find the company website? Can I find what they do? I would love just throwing all that stuff in there. I want that in my read summary. If I can put

a little blurb next to the activity or next to the company." There are 18 instances of this code among six of the readers.

Finally, the photo-narrative activity provided for a highly reflective opportunity for the reviewers. They noted how useful the activity was in helping them reflect on their own review processes. P4 remarked, *"I think it was very useful to realize how much work we're looking at and how much we're potentially talking about for just one application. And each of us we read over a thousand applications a year...It's a lot of information for one application. We all realized it but when you actually take pictures of it and lay it out on paper and look at it from more of a macro-level view, it really is just interesting."* Several of them suggested that this activity would perhaps be useful for training future reviewers. P1 said, *"I think this is honestly a good exercise to consider when we are training our new admissions counselors...Let us say this is a committee for early action, show us what you guys are thinking about...We speak a certain language and we may be like, why did you highlight this one sentence?"*

Discussion of Workshop #2. The themes identified around the photo-narrative activity proved exciting and useful to the entire team of technologists and reviewers. This activity allowed reviewers to reflect on their own practices and provide the entire team with ideas on how to proceed in support of more effective and efficient committee meetings. Two key design observations emerged from the activity.

Tangible evidence representation lends itself to storytelling with rationale. Often, tools such as .pdf readers support the embedding of notes within documents and provide a listing of those notes — much like the Adobe Acrobat .pdf viewer. The existing tool that our reviewer participants use for managing review documents works similarly, allowing reviewers to embed notes throughout the various parts of the application. While this may support recall, allowing the reader to quickly browse to find and remember important points when preparing for the committee meeting, it does not adequately support storytelling. The photo narratives generated during the activity indicate that given the opportunity, reviewers generate stories that are temporal and include linking and grouping of relevant evidence, and that these stories aligned with several genres from the visual narratives literature (RQ2).

Established rubrics dominate story structure. One of the most common codes from the photo-narrative activity is that the stories were structured around elements of the *Decision Sheet* as seen in the way that the readers grouped and ordered their evidence on their narratives. For example, P4 clustered his evidence by *Decision Sheet* categories and ordered them the way they would appear in the *Decision Sheet* (i.e. demographics first, then academics, then non-academics). Even other readers who did not have such an explicit structure tended to order academic information before non-academic information as is currently done in their *Decision Sheet*. Recall that the *Decision Sheet* is essentially a summary for the rubric used by reviewers. This rubric is the cornerstone of the decision-making process and is front-and-center during the committee review discussions.

Communication of evidence weighting/sentiment is important. Knowing how important a piece of evidence is to a reviewer, and whether it's positive or negative, is important for the others in the committee to calibrate and understand the leaning of a reviewer. The weighting is not necessarily numerical, but can be a binary weighting — positive or negative. By making this information visible, a reviewer provides an overview of their "lean" for the applicant, framing the discussion around those positive and negative pieces of evidence.

Stories support training. The photo-narrative activity carried out with our reviewers had unintended yet interesting outcomes in that it illuminated opportunities for training. Constructing stories visually to show a reviewer's argument for a student can be an invaluable training tool for new reviewers. By integrating these capabilities into committee review software, training could be

ongoing and happen in each meeting. It would also serve to help calibrate reviewers — another aspect mentioned throughout our workshop studies.

In summary, stories were effectively told with this evidence and they were organized around the *Decision Sheet* that loomed large in the thinking of the reviewers. Additionally, these stories served to help train and calibrate new reviewers. These insights led to the final PD workshop in which the team explored design ideas for harnessing evidence in support of storytelling *within* the *Decision Sheet* framework.

3.4 Workshop 3: Decision Sheet Redesign

Review processes often revolve around a rubric that describes the criteria of interest. As described in [38] and from our observations in Workshop #2, holistic review typically employs a *Decision Sheet* of some sort where information critical to the reviewers' recommendations (e.g., notes, ratings) are recorded. This *Decision Sheet* serves as somewhat of a "rubric" in that it is structured by the criteria over which an applicant is evaluated. In the case of our admissions team, this *Decision Sheet* was the center of the discussion during committee and drove the process. As we observed in Workshop #2, this *Decision Sheet* also guided storytelling, even when not present in the photo-narrative exercise. In this final workshop, we sought to explore the design of the *Decision Sheet* to identify opportunities for improvement with a particular eye for supporting evidence-based storytelling. Our team discussed opportunities for incorporating evidence and story more directly into the *Decision Sheet* with particular attention to how to leverage story structures from [33] and those found in Workshop #2, within the confines of the *Decision Sheet*.

We reviewed the existing *Decision Sheet* together. The current sheet is distributed over two "pages" thus making it difficult to view the entire rubric at once. Our team expressed concerns over the poor use of space, a lack of space for including their notes, content organization that inadvertently hid important information, and a lack of context, among others. We also discussed opportunities for integrating evidence more directly with the *Decision Sheet*. By definition, the *Decision Sheet* already uses a clustering visual narrative tactic where related information is grouped together. We aimed to work within and/or alongside this structure. We developed paper prototypes as well as an interactive digital prototype to explore our ideas and iterated on those ideas always with feedback and input from our admissions team.

Design Iteration Findings. Design #1 as shown in Figure 3 integrated a slideshow model into the existing *Decision Sheet*. In this design, the lead reader would step through the "story" using the buttons to advance from one key story point to the next where each key point would contain some form of evidence such as a text highlight or snapshot. For each key point, the red box would highlight the facet of the *Decision Sheet* with which that key point was associated. Additional relevant reader notes and comments, as well as the evidence, would be shown upon demand in a pop-up window (i.e., messaging tactic) when the user clicked on the box. While readers appreciated the ability to represent and move through the *Decision Sheet* as a story, they were concerned with the rigidity and potential slowness of this model. P7 worried that it would be too constricting, saying *"I think it should be left up to the storyteller to decide what to highlight. Each student/story is different."* P6 remarked, *"I think this might slow us down as there might be too many areas of interests. We want staff to be precise — I think this opens up too many unnecessary conversations."*

Design #2 explored a more explicit representation of the evidence. In this design shown in Figure 4, the colored dots represent important points of evidence (i.e., highlighting visual narrative tactic) that the lead reader wants to bring up during the committee meeting, and the color represents positive (green), negative (red) and neutral (yellow) sentiment about those points of evidence. A reader telling the student's story during a committee meeting can move the mouse pointer from one point of evidence to the next on the *Decision Sheet* in any desired order. When the mouse cursor

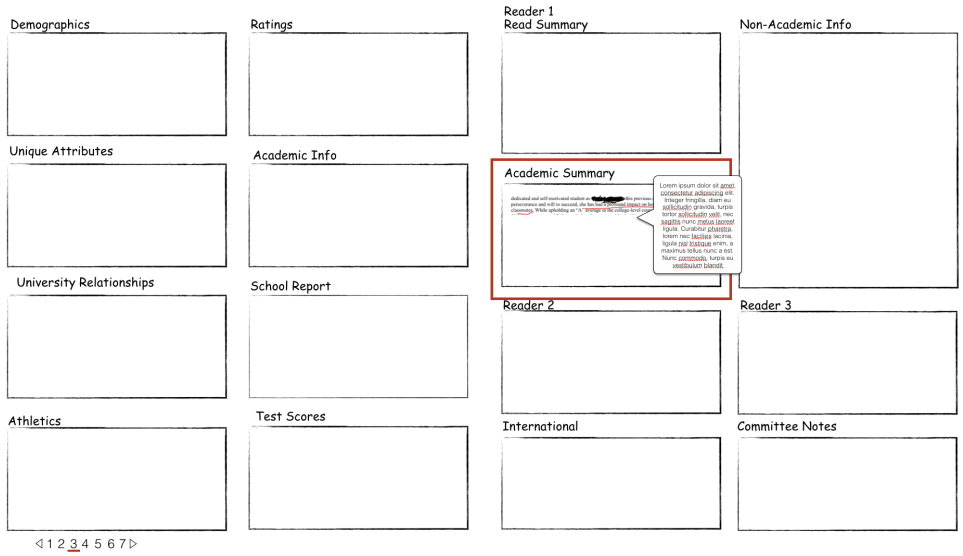

Fig. 3. Mockup of a design that emphasizes the slide-show approach. The slides are numbered 1-7. In a typical slideshow, the user would step through the slides and be shown one slide at a time. In this design, the slide "window" moves throughout the *Decision Sheet* to highlight the facet of interest. Note that in these mockups, each empty box represents a facet of the *Decision Sheet* (or rubric) and would contain summary information for that facet in text or chart form.

Fig. 4. Design #2 that emphasizes evidence markers, but no specific order.

is over a point of evidence, additional information (notes, evidence snapshot, etc.) is highlighted and shown in a pop-up window or tool tip. In Figure 4, the reader is hovering over an evidence point in the Non-Academic Info notes section, and additional information is shown in the grey popup box.

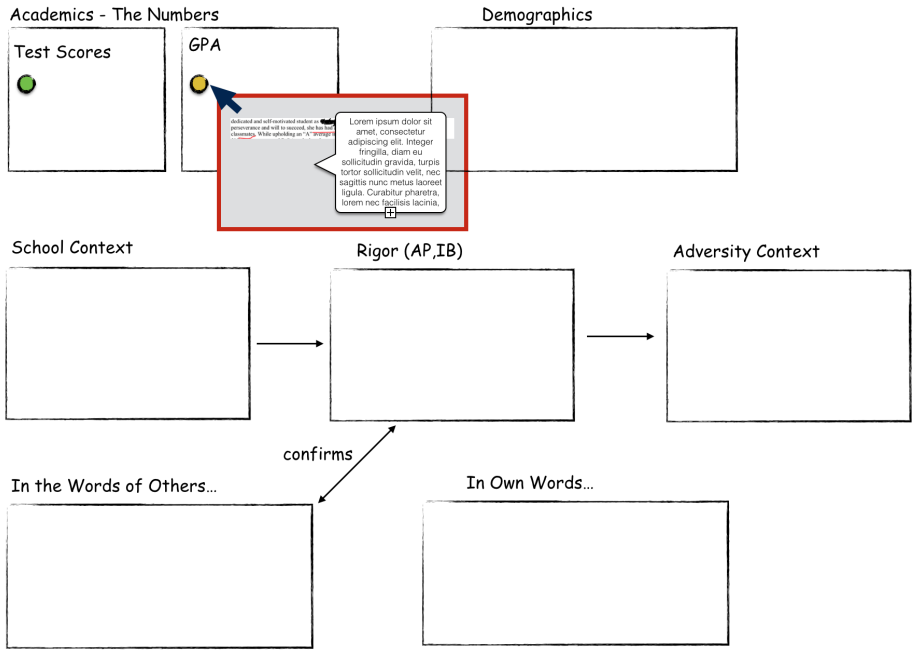

Fig. 5. Design #3 allows complete freedom in layout and order along with inclusion of the evidence markers.

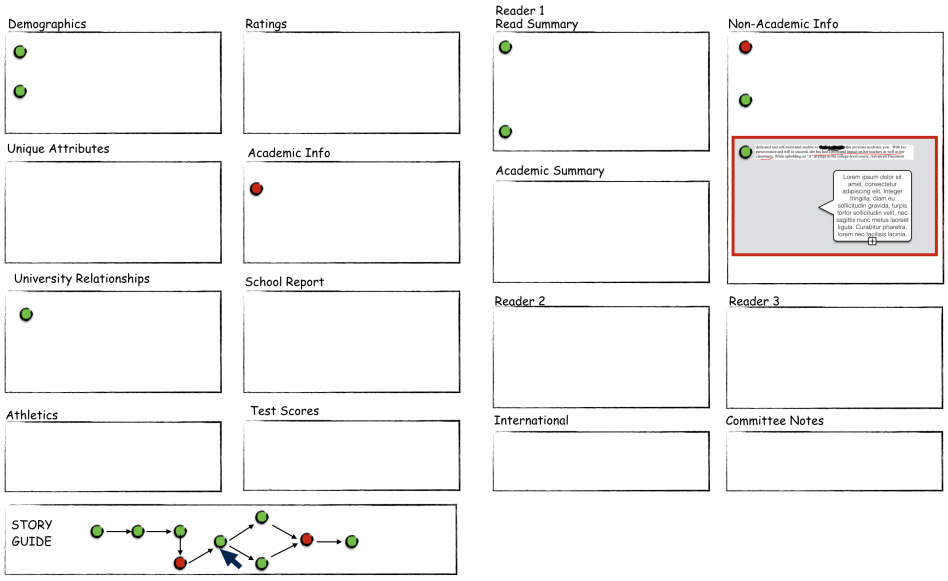

Fig. 6. This iterative refinement combines aspects of multiple designs, integrating evidence markers while providing for customization of the story within context of the standardized *Decision Sheet*.

Readers liked that the evidence dots provided an overview of the lead reader's view. P4 noted, *"I like how you can quickly see which way someone may be leaning in regards to their final recommendation due to the corresponding color coding."* They also appreciated that the evidence dots helped them locate important information for the story, *"I like the fact that you can bring up your snapshot from the app without re-opening the app again and searching for that information"* [P4], however, they were not sure how this approach might affect their overall time to review an application. Three reviewers specifically noted that an advantage of the design is that it would make their process easier since they can directly focus on the specific areas were of interest to them for this particular applicant. P4 remarked that *"This design would more than likely allow us to jump immediately to certain areas of support or concern which would assist us with focusing on the most important details during our discussion."* For P1, it was useful for cases where she wasn't quite sure about an applicant, *"Sometimes, I'm not sure what to do with a student but it'll be useful to say — I know I want to draw attention to these specific points (evidence)"*. There was disagreement, however, on whether there should be an "order" to the progression through the evidence dots.

A third design (Figure 5) also used the the colored evidence dots, however, the *Decision Sheet* elements in this design are arranged in a completely custom fashion by the reviewer using arrows as a visual narrative structure to indicate sequencing. The idea was to examine how flexibility and creativity were viewed by the readers. Readers welcomed the ability to be creative in organizing their story in a freeform flowchart-inspired manner. P4 stated, *"Having the ability to move the various parts of the story-line and evidence around to organize the 'story' itself would be helpful"*. However, they expressed concerns that it might give the readers too much freedom and could cause confusion because it would be difficult for reviewers to locate typical elements of the *Decision Sheet*. P3 said, *"This will be easy for the storyteller to read, but may take some time for other review committee members to find the information they are looking for."*

A hybrid design, shown in Figure 6, retains the evidence markers for evidence (red and green dots), and provides a separate window that allows the reviewer to organize that evidence into a particular free-form story (i.e., flowchart genre). As the reviewer hovers over dots in that view, the red box highlights the actual elements of the *Decision Sheet* where that evidence exists and allows for further exploration of details. In this design, reviewers can custom organize evidence as they see fit into a structured "story" employing arrows and clusters as narrative tactics and use that organization to guide them through the standard *Decision Sheet* layout.

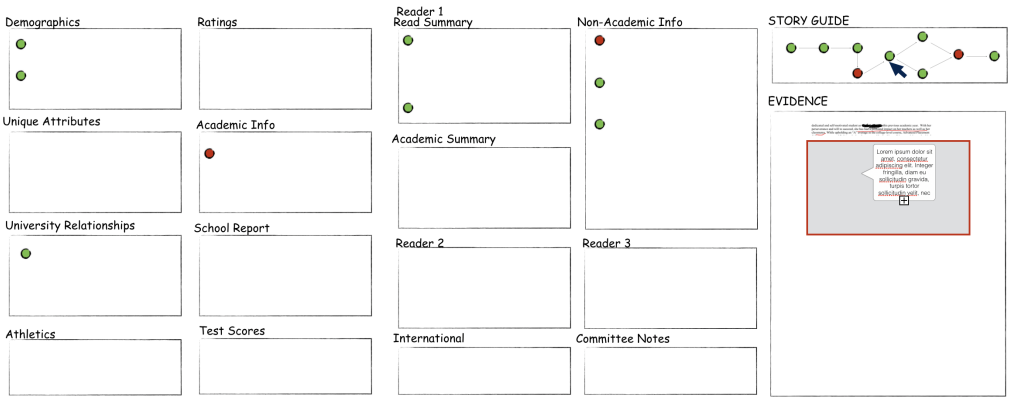

Fig. 7. This iteration builds on previous designs to build a custom "storyline", but shows the evidence in a separate box to declutter the *Decision Sheet* contents.

All the readers liked the story guide, but some were concerned about the additional time overhead. P1 said, *"I think the story guide will help committee be more efficient, especially [for] those who have not prepared their files prior to entering a committee room. My concern is for those who do take time to prepare for committee — does the story guide add more prep time?"* Other readers commented on how it would fit into the current process, especially in regards to the roles of committee members. P4 remarked, *"I believe there is the potential here to decrease the time that we would use for the initial 'read' of the application as well as the preparation for and execution of review committee itself. I can still see the value of the 'roles' for the review committee discussions, though. It is difficult for me to see us not having the leader of the review committee (moving the files, decisioning the files, etc.), the presenter (territory manager), and the 'members' weighing in."*

A further iteration on this design seen in Figure 7, has a customized sequential storyline and a separate pane for relevant evidence. By having a dedicated pane for the evidence, the content that is normally seen in the *Decision Sheet* will not be cluttered by or confused with the evidence snapshots.

To summarize, in this final workshop, our aim was to explore concepts for a *Decision Sheet* augmented with visual evidence. We created multiple iterations on the concepts in an effort to find a balance between the flexibility and standardization that are necessary in crafting an applicant's story within the confines of the *Decision Sheet*. We also wanted to provide opportunities for readers to showcase their leanings as well as the evidence that landed them there.

4 DISCUSSION

The analysis of storytelling in review processes in general, and holistic admissions in particular, is absent in prior work. This is not a surprise given the sensitive nature of the topic and the time constraints on reviewers in real admissions processes. Here, we have explored and characterized a particularly important aspect of holistic review — the committee review, in which readers discuss an applicant, argue for their acceptance, and collaborate to come to a recommendation.

The workshop structure of our participatory design activities proved appropriate for our particular situation, where it was important to make focused progress in exploring RQ1-RQ3 in relatively short meetings due to the time constraints of our PD team. To recap, the first workshop was designed to characterize the challenges and opportunities of the committee review aspect of holistic admissions review. We found, among other things, that evidence collection and memory challenges directly affect the story that readers tell during their committee review meetings. Workshop #2 provided an opportunity to explore storytelling structure through a photo-narrative activity. It became clear that stories have somewhat of a structure and that structure was influenced largely by the *Decision Sheet*, a rubric for summarizing an applicant (RQ1). We confirmed the importance of the *Decision Sheet* structure and of directly connecting evidence and commentary to it. We further categorized photo-narrative stories into a set of visual narrative genres to demonstrate opportunities for supporting different styles of storytelling (RQ2). Finally, in Workshop #3, the team iterated over multiple design ideas that sought to redesign the *Decision Sheet* with an eye for integrating evidence directly and supporting evidence recall during storytelling (RQ3).

While still in the early phases of design for this domain, we reflect on our experience in co-designing with admissions officers, and we discuss design inspirations drawn from our participatory design activities. As articulated by Bratteteig et al. [9], a PD result can be many things, including an understanding of design innovations to empower participants to improve their process, as well as a reflective look at the process itself and the opportunities and challenges it presents. We conclude this discussion by drawing parallels of holistic admissions to other review processes and how and where our findings may apply.

4.1 Review Processes are Sensemaking Processes

Committee review stories are the culmination of a broader collaborative sensemaking process [38] that must be considered in the design of technological support tools for reviewers who serve as the analysts whose job it is to review the evidence and make recommendations on a course of action (e.g., accept or reject). One might assume that storytelling is confined to the committee meeting discussion. However, while the application data are prepared and clean and all the data exists at the time of review, the readers must still ferret out the important elements of the application and connect the dots between the elements. Thus, the other key elements of sensemaking that revolve around identifying and marshalling evidence and subsequently organizing that evidence into a hypothesis are key to design thinking for innovative solutions. Findings from visual and collaborative sensemaking work [5, 20], while not focused on evidence gathering and storytelling, may help point towards designing possible solutions.

We suggest that proper treatment of evidence, both collection and representation, are of critical importance in supporting review processes such as that in holistic admissions review.

Storytelling begins with evidence capture. As we discovered in Workshop #1, committee review produces significant cognitive strain on readers who read tens of applications a day and must come to committee review meetings (sometimes several days later) prepared to represent a student that they would like to argue for in favor of admission. While many application and document management tools support highlighting and note-taking, they generally keep this information attached to and distributed throughout the application and/or aggregated in a separate "notes" window or pane. This is useful for revisiting the notes and perhaps recalling interesting and important aspects of the application, however, they lack a connection to a rubric and thus do not serve well as rationale for a recommendation. Furthermore, rereading an application is a time-consuming process that readers seek to avoid when preparing for committee review — often relying on their own notes (e.g., handwritten and/or embedded in the application).

We call for designers to rethink interaction design for rubric-based review situations. In these cases, the rubric (i.e., the *Decision Sheet* for our admissions reviewers), is likely the centerpiece of the committee review. By associating evidence directly with aspects of the rubric, designers help reviewers capture their rationale for a recommendation. Our studies indicate that design in support of review should not simply support note-taking and highlighting or aggregation of notes, but direct connection of that "evidence" to the rubric. Inspired by the importance of visual histories in supporting recall after task interruptions [32], we recommend the exploration of the idea of allowing reviewers to capture visual histories of relevant evidence [32] through visual snapshots. This will help readers reconstruct their rationale from the snapshot moment, thus supporting preparation for storytelling during committee review [32, 43]. Our photo-narrative activity further supports this approach showing that our domain experts were able to tell compelling applicant stories using evidence snapshots.

Given the time-sensitive nature of application review, this approach to evidence capture might raise concerns over efficiency. In fact, while our initial goal was to explore storytelling with evidence under the assumption that the evidence could be easily captured or extracted, we quickly found that it was difficult to have these kinds of hypothetical conversations with the admissions officers. They simply struggled to explore evidence representation without understanding how that evidence would be extracted.

In response to this concern, we carried out a brief evidence capture experiment. We implemented a simple snapshot tool that could be executed in a browser to demonstrate to the readers how this process could work in practice. This exercise proved fruitful not just to put our stakeholders at ease, but to generate new ideas. In this experiment, reviewers were asked to capture a relevant aspect of

an academic portion of an application using a custom built screen capture tool that allowed the user to drag a box over the area of interest to create the snapshot. Upon capture, a popup window would open and ask the user to rate that piece of evidence as positive(green), negative(red), or neutral(yellow) by clicking on the color-coded radiobox that best described the evidence (See the example of evidence capture flow included in the supplemental materials). The average rating on a 7-Point Likert scale (1-very unacceptable, 4-neutral, 7-very acceptable) was 4.5. In other words, they were not concerned that it would negatively impact the time to review an application.

Storytelling requires authoring. Despite our efforts to focus on the storytelling aspect of committee review, our research team found that as a participatory design team with our admissions reviewers, we could not avoid questions regarding *authoring*. As noted above, reviewers were anxious to understand how the evidence would be captured efficiently. Likewise, throughout Workshop #3 activities, our reviewers were unclear as to *how* the evidence would be associated with parts of the *Decision Sheet* and/or ordered in the designs that represented the evidence in some sort of sequence (e.g., Figures 6 and 7). We found that interactive prototypes to demonstrate these aspects proved useful. The evidence capture demo described above not only allowed the reviewers to comment on the acceptability of such an interaction, but also inspired new ideas.

For example, consider that holistic admissions reviewers in our case develop domain specific terms and shorthand for describing aspects of an application. For example, a student might be described as "spunky" and all reviewers would have an understanding of that term based on previous uses of it in context. Reviewers suggested that an evidence-capture snapshot tool should allow them to not only rate evidence, but *tag* the evidence with their specific language or codes. Such tags could then be exported to the *Decision Sheet* and associated directly with the evidence. We encourage designers to consider support of evidence capture as well as tagging in holistic review situations. A capture tool, for example, might be configured with a list of tags commonly used and allow reviewers to simply click a checkbox to select that tag and associate it with a snapshot.

While we did not explore prototypes for authoring story structures, we see promise in multiple approaches. For rubric-driven domains such as ours, we encourage exploration of storytelling that can be directly integrated with the rubric structure. While the rubric provides categories over which evidence may be associated, it does not provide sequencing. Our reviewers fell on both sides of the sequencing argument. Some reviewers wanted complete freedom to organize the evidence into a specific sequence to tell their story in a creative way. Others preferred a more standard sequencing, perhaps driven by the *Decision Sheet*, and that would create a more consistent story and require less overhead (and would perhaps be less biased by reviewer storytelling skill). Any solution would have to be efficient — at least as efficient as reviewers are when they currently prepare for committee by reviewing the application and preparing story notes.

Storytelling requires representation and narration. Our original focus was on the narration aspect of the committee meeting story, the final piece of the puzzle in presenting evidence in a story. That is, we sought to explore how to support the reviewers in verbally presenting that applicant during committee review. As noted above, we quickly found that we had to also consider how to capture evidence and author story elements. Nonetheless, our workshops resulted in several interesting findings regarding how to represent evidence in support of that narrative.

First, the evidence representation provides an overview of the reader's take on the student. Perhaps, for example, all evidence is positive, but not exciting. On the other hand, a student can have several positives along with a serious negative. We call for designers to visually represent evidence and its valence in support of storytelling in collaborative situations where the story is being conveyed to a team that will jointly make a recommendation. This approach to providing a

quick overview aligns well with the information visualization mantra of Shneiderman, "Overview first...then details on demand" [34]. In our prototypes, details were provided by clicking on specific evidence dots to recall the actual evidence.

We also urge designers to explore how to represent the "importance" of each piece of evidence. Not all evidence is considered equal; three positives, for example, might be balanced out by a very serious negative. Additionally, representation of the importance of such evidence has been shown to be important in combating confirmation bias [11] in intelligence analysis situations.

Finally, as illustrated in the results of Workshop #2, while structured by the *Decision Sheet* rubric, our reviewers still each had somewhat unique ways of telling stories. While we, as a team, chose to focus on remaining within the structure of the *Decision Sheet* in this particular design case, we encourage designers to explore other narrative options guided by our genre alignment findings shown in Table 1.

Visual storytelling supports training and calibration. One side-effect of Workshop #2 is that the team was able to reflect on their process from a new perspective. The photo-narrative exercise led to interesting observations about what they do and how they learn to do it.

Less-experienced reviewers stand to learn a great deal from more experienced reviewers when their reasoning is exposed and highlighted. They can benefit from seeing what evidence was considered and how it is weighted — whether positively or negatively. Visual representation of evidence and story can help less-experienced reviewers learn to organize evidence into a coherent and compelling narrative. Thus, we suggest that designers consider the role of visual representations of evidence for training and calibration within review processes, perhaps drawing knowledge from previous work in information sharing [3, 7, 21].

4.2 Cognitive Bias Mitigation in Storytelling

There are also opportunities for cognitive biases to creep into the process [38]. In particular, during committee, reviewers may be affected by "availability bias" due to retrievability of instances in that their initial reading of the documents possibly happened many days before and their ability to recall all of the relevant evidence for this argument depends on their memory, note-taking, and any other cognitive supports employed for reviewing applications. Reviewers may also be affected by the narrative fallacy where there is a tendency to present coherent, "clean" stories when, in fact, the situation may be quite nuanced [19]. Finally, reviewers may succumb to *confirmation bias* where they present the information about an applicant that tends to support their initial judgments. All of these cognitive biases may directly manifest themselves in the "storytelling" aspect of review meetings where a leader must present a case. We have also come to realize somewhat conflicting perspectives regarding parts of the process that suggest other potential biases. In particular, while some reviewers want the freedom to tell their stories, others see value in the consistency of stories. What is best for avoiding cognitive bias and recommending the most intriguing students when many of them meet or even exceed the academic requirements? How do we level new vs. experienced readers?

As noted in Workshop #1 results, committee review is a highly collaborative process where the goal is for the team to make the best recommendations. That is, even though a single reviewer will represent their particular set of applicants during committee review, everyone's goal is to make the right recommendations. To this end, each person must play their role and while the lead reviewer's goal is to paint an honest picture for that student, the others must confirm their points and also attempt to poke holes in the story. Having all important evidence present and accessible from the *Decision Sheet* helps all roles and provides a level of transparency even in the face of varying storytelling abilities and inevitable cognitive biases.

4.3 Parallels To Other Review Processes

Although this paper focuses on holistic admissions review, we suspect our findings are applicable to other review processes. In particular, the *mechanics* of the committee meetings of holistic admissions review and the *mechanics* of the common committee- or panel-based peer review process are quite similar in many respects. In particular, they all involve three key elements of interest in this paper:

- An initial individual reading/review (e.g., applications, papers, proposals)
- A subsequent committee meeting or review panel
- An assigned "lead" for the discussion

One can also draw other parallels between the holistic admissions review and these other review processes. Consider, for example, that holistic review committee meetings are largely "explanatory" rather than "exploratory" in nature where the lead reviewer explains why a student should (or should not) be admitted, often pointing to evidence found in the application. While other members of the committee may briefly "explore" the application during committee as the lead reviewer presents the applicant, there is limited time for exploration. These roles (e.g., lead reviewer and secondary reviewers) are similar in peer-review panels and the time constraints are similarly present; thus, the discussions are largely explanatory where the lead reviewer's opinion and that of anyone else who has carefully read the proposal carries weight. Likewise, in all of these cases, final decisions are made by a "chair" or "director" who may need to fall back on reviewer notes and recommendation rationale to make final decisions.

On the other hand, we also understand that each domain contains very different data and knowledge during the review process (e.g. blind review). While all of these differences certainly have an effect on the overall process, the need to tell a compelling story based on evidence is present and further work is needed to see how our findings extend to other review domains.

5 DESIGN CONSIDERATIONS AND FUTURE WORK

In Section 4, we presented numerous opportunities for innovation and exploration in the design of software to support holistic review processes for college admissions and possibly other similar holistic review domains. These innovations stand to help reviewers and ultimately lead to better outcomes. In summary, we call for designers to consider the following when creating solutions for holistic review domains:

- Give the rubric and evidence high priority and allow for direct association between them.
- Consider visual histories to support evidence capture and recall.
- Consider support of evidence capture, weighting, and tagging during review.
- Explore ways to integrate the rubric in storytelling.
- Visually represent captured evidence, its valence, and its weighting in order to provide an overview of the reviewer's "leaning."
- Explore existing visual narrative genres as well as novel hybrids for telling review stories effectively.
- Consider training and calibration as a potential side-effect of effective designs.
- Consider bias mitigation strategies at all times.

In addition to these considerations, much future work is needed, in particular, to understand how to mitigate the various cognitive biases that may occur, to determine how our results apply to other review domains, and to determine how to support holistic review processes more effectively with visual representations of the data sources.

Finally, while we focused on visual representations of evidence and rationale rather than visualization design of the applicant data itself, there are still many opportunities for exploration in future work. In particular, we found that the current text-based representation of the *Decision Sheet*

is lacking in several aspects, especially concerning the representation of an applicant's data in the context of other applicants (e.g., from the same high school or from the current admitted class). Showing a student in the context of many (tens or hundreds) of other students could prove valuable and would likely require proper visualization and interaction design to integrate into their existing *Decision Sheet*.

6 CONCLUSION

To conclude, we expanded on prior field work [37] to explore more deeply the role of storytelling in the holistic admissions domain. Through three participatory design workshops, we discovered commonalities and differences in how stories are composed and told and characterized these stories using categorizations from existing literature on visual narratives. Through co-design iterations with a college admissions team, we explored possible *Decision Sheet* redesigns for incorporating storytelling and learned about how they could impact the holistic review process. Finally, we provide design suggestions and future research directions regarding storytelling and evidence capture in review processes.

ACKNOWLEDGMENTS

We would like to thank our team of holistic admissions review officers for contributing their time, expertise, and creativity to this project. We would also like to thank Poorna Talkad Sukumar for fruitful conversations regarding this work. This material is based upon work supported by the National Science Foundation under Grant No. 1816620.

REFERENCES

[1] 2002. Best Practices in Admissions Decisions. *College Entrance Examination Board* (2002).
[2] 2019. College Admissions Scandal. https://www.nytimes.com/news-event/college-admissions-scandal. Accessed: 2019-09-13.
[3] Mark S Ackerman, Juri Dachtera, Volkmar Pipek, and Volker Wulf. 2013. Sharing knowledge and expertise: The CSCW view of knowledge management. *Computer Supported Cooperative Work (CSCW)* 22, 4-6 (2013), 531–573.
[4] Sergio Alonso, Enrique Herrera-Viedma, Francisco Javier Cabrerizo, Carlos Porcel, and Antonio Gabriel López-Herrera. 2007. Using visualization tools to guide consensus in group decision making. In *International Workshop on Fuzzy Logic and Applications*. Springer, 77–85.
[5] Jeff Baker, Donald Jones, and Jim Burkman. 2009. Using visual representations of data to enhance sensemaking in data exploration tasks. *Journal of the Association for Information Systems* 10, 7 (2009), 2.
[6] Michael N Bastedo, Joseph E Howard, and Allyson Flaster. 2016. Holistic admissions after affirmative action: Does "maximizing" the high school curriculum matter? *Educational Evaluation and Policy Analysis* 38, 2 (2016), 389–409.
[7] Alexander Boden, Frank Rosswog, Gunnar Stevens, and Volker Wulf. 2014. Articulation spaces: bridging the gap between formal and informal coordination. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 1120–1130.
[8] William G Bowen and Derek Bok. 1999. *The shape of the river: Long-term consequences of considering race in college and university admissions*. Princeton University Press Princeton, NJ.
[9] Tone Bratteteig and Ina Wagner. 2016. What is a participatory design result?. In *Proceedings of the 14th Participatory Design Conference: Full papers-Volume 1*. ACM, 141–150.
[10] Wayne J Camara and Amy Elizabeth Schmidt. 1999. Group Differences in Standardized Testing and Social Stratification. Report No. 99-5. *College Entrance Examination Board* (1999).
[11] Maia B Cook and Harvey S Smallman. 2008. Human factors of the confirmation bias in intelligence analysis: Decision support from graphical evidence landscapes. *Human Factors* 50, 5 (2008), 745–754.
[12] Sarah E Drew, Rony E Duncan, and Susan M Sawyer. 2010. Visual storytelling: A beneficial but challenging method for health research with young people. *Qualitative health research* 20, 12 (2010), 1677–1688.
[13] Daniel Golden. 2007. *The Price of Admission: How America's Ruling Class Buys Its Way into Elite Colleges-and Who Gets Left Outside the Gates*. Broadway Books.
[14] James Hollan, Edwin Hutchins, and David Kirsh. 2000. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)* 7, 2 (2000), 174–196.
[15] Karen Holtzblatt and Hugh Beyer. 1997. *Contextual design: defining customer-centered systems*. Elsevier.

[16] M Shahriar Hossain, Patrick Butler, Arnold P Boedihardjo, and Naren Ramakrishnan. 2012. Storytelling in entity networks to support intelligence analysts. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1375–1383.

[17] Joshua E Introne. 2009. Supporting group decisions by mediating deliberation to improve information pooling. In *Proceedings of the ACM 2009 international conference on Supporting group work*. ACM, 189–198.

[18] Nipat Jongsawat and Wichian Premchaiswadi. 2009. Group awareness information in web-based group decision support system. In *2009 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, 370–375.

[19] Daniel Kahneman. 2011. *Thinking, fast and slow*. Macmillan.

[20] Paul E Keel. 2007. EWall: A visual analytics environment for collaborative sense-making. *Information Visualization* 6, 1 (2007), 48–63.

[21] Benedikt Ley, Thomas Ludwig, Volkmar Pipek, Dave Randall, Christian Reuter, and Torben Wiedenhoefer. 2014. Information and expertise sharing in inter-organizational crisis management. *Computer Supported Cooperative Work (CSCW)* 23, 4-6 (2014), 347–387.

[22] Narges Mahyar, Weichen Liu, Sijia Xiao, Jacob Browne, Ming Yang, and Steven P Dow. 2017. ConsesnsUs: Visualizing Points of Disagreement for Multi-Criteria Collaborative Decision Making. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 17–20.

[23] Narges Mahyar and Melanie Tory. 2014. Supporting communication and coordination in collaborative sensemaking. *IEEE transactions on visualization and computer graphics* 20, 12 (2014), 1633–1642.

[24] Helena M Mentis, Paula M Bach, Blaine Hoffman, Mary Beth Rosson, and John M Carroll. 2009. Development of decision rationale in complex group decision making. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1341–1350.

[25] Michael J Muller. 2009. Participatory design: the third space in HCI. In *Human-computer interaction*. CRC press, 181–202.

[26] Iván Palomares and Luis Martínez. 2014. Low-dimensional visualization of experts' preferences in urgent group decision making under uncertainty. *Procedia Computer Science* 29 (2014), 2090–2101.

[27] Iván Palomares, Luis Martínez, and Francisco Herrera. 2014. MENTOR: A graphical monitoring tool of preferences evolution in large-scale group decision making. *Knowledge-Based Systems* 58 (2014), 66–74.

[28] Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, Vol. 5. McLean, VA, USA, 2–4.

[29] Katie Reilly. 2018 (accessed August 12, 2019). *A Lawsuit by Asian-American Students Against Harvard Could End Affirmative Action as We Know It*. https://time.com/5425147/harvard-affirmative-action-trial-asian-american-students/

[30] Michael Ross and Fiore Sicoly. 1979. Egocentric biases in availability and attribution. *Journal of personality and social psychology* 37, 3 (1979), 322.

[31] Adam Rule, Aurélien Tabard, and Jim Hollan. 2017. Using visual histories to reconstruct the mental context of suspended activities. *Human–Computer Interaction* 32, 5-6 (2017), 511–558.

[32] Adam Rule, Aurélien Tabard, and Jim Hollan. 2017. Using visual histories to reconstruct the mental context of suspended activities. *Human–Computer Interaction* 32, 5-6 (2017), 511–558.

[33] Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 1139–1148.

[34] Ben Shneiderman. 1996. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE symposium on visual languages*. IEEE, 336–343.

[35] Jacques Steinberg. 2003. *The gatekeepers: Inside the admissions process of a premier college*. Penguin.

[36] Anselm L Strauss. 1987. *Qualitative analysis for social scientists*. Cambridge university press.

[37] Poorna Talkad Sukumar and Ronald Metoyer. 2018. A Visualization Approach to Addressing Reviewer Bias in Holistic College Admissions. In *Cognitive Biases in Visualizations*. Springer, 161–175.

[38] Poorna Talkad Sukumar, Ronald Metoyer, and Shuai He. 2018. Making a Pecan Pie: Understanding and Supporting The Holistic Review Process in Admissions. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 169.

[39] Jan M Van Bruggen, Henny PA Boshuizen, and Paul A Kirschner. 2003. A cognitive framework for cooperative problem solving with argument visualization. In *Visualizing argumentation*. Springer, 25–47.

[40] Susan W van den Braak, H van Oostendorp, Henry Prakken, and Gerard AW Vreeswijk. 2008. A critical review of argument visualization tools: Do users become better reasoners?. In *Workshop notes of the ECAI-06 workshop on computational models of natural argument*. 67–75.

[41] Douglas Walton. 2009. Argument visualization tools for corroborative evidence. In *Proc. of the 2nd International Conference on Evidence Law and Forensic Science*. 32–49.

[42] Lu Xiao and Richelle L Witherspoon. 2015. Information sharing as story construction in group decision making. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community*.

American Society for Information Science, 63.
[43] Lingxue Yang. 2018. *UX design for memory supplementation to support problem-solving tasks in analytic applications.* Ph.D. Dissertation.
[44] Qiyu Zhi, Suwen Lin, Poorna Talkad Sukumar, and Ronald Metoyer. 2019. GameViews: Understanding and Supporting Data-driven Sports Storytelling. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* ACM, 269.

Received January 2020; accepted March 2020