Pathways: Exploring Academic Interests with Historical Course Enrollment Records

Youjie Chen
Cornell University
Ithaca, NY, USA
yc2669@cornell.edu

Annie Fu
Cornell University
Ithaca, NY, USA
af397@cornell.edu

Jennifer Jia-Ling Lee
Cornell University
Ithaca, NY, USA
jjl296@cornell.edu

Ian Wilkie Tomasik
Cornell University
Ithaca, NY, USA
iw68@cornell.edu

René F. Kizilcec
Cornell University
Ithaca, NY, USA
kizilcec@cornell.edu

ABSTRACT

Students are encouraged to explore their interests during college to stimulate intellectual growth and prepare for a dynamic labor market. However, interest exploration is entangled with the fateful process of choosing courses for enrollment, and most institutions offer limited tools to help students choose. We propose *Pathways*, an interactive course information retrieval tool that facilitates interest exploration and course discovery with a diverse pool of historical course enrollment records. The tool visualizes sequences of course enrollments as “academic pathways” to grant students unprecedented insights into the academic choices of prior students. We share our design process, including a formative study on need analysis, the UX and algorithm design, and an evaluation study. We find that *Pathways* supports students in finding courses that both match their interests and expose them to new ideas. We discuss directions for future work on how interest exploration can be promoted at scale and on how to utilize historical course enrollment data through visualization.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; • **Applied computing → Education**.

KEYWORDS

Interest, Course Decision, Data Storytelling, Higher Education

ACM Reference Format:
Youjie Chen, Annie Fu, Jennifer Jia-Ling Lee, Ian Wilkie Tomasik, and René F. Kizilcec. 2022. Pathways: Exploring Academic Interests with Historical Course Enrollment Records. In *Proceedings of the Ninth ACM Conference on Learning @ Scale (L@S '22), June 1–3, 2022, New York City, NY, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3491140.3528270

L@S '22, June 1–3, 2022, New York City, NY, USA.

ACM ISBN 978-1-4503-9158-0/22/06…$15.00
https://doi.org/10.1145/3491140.3528270

1 INTRODUCTION

Each year thousands of jobs disappear and new industries and occupations emerge [15]. College students are starting careers with rapidly evolving demands due to the transformation of entire industries by digitization and automation [39, 47]. Three out of four US college graduates work in a job that is not even related to what they studied in college [1]. These trends make it more important than ever for college students to actively explore their interests to gain awareness of potential industries to work in and acquire a range of knowledge and skills to adapt in an increasingly complex and rapidly-changing world [30]. To complement a structured curriculum with required courses, the elective choice system in US higher education grants students freedom to explore a variety of courses during their degree program. The resulting sequence of course enrollments across terms, their *academic pathway*, is a relatively unique reflection of their formal academic experience [4, 5, 19].

Exploring and developing an academic pathway is a *constructivist* learning process and distinct from acquiring skills and knowledge in a specific domain [26, 73]. Transcending the learning that goes on in classrooms, students continuously explore, construct, and reflect on their learning path, as they will continue to do after they graduate. Moreover, pathway exploration and development emphasizes that knowledge is constructed from one's own experience. Students iteratively develop pathways by trying new things themselves, as opposed to being instructed by an expert about the right path. Although many colleges and universities encourage constructivist approaches to learning, most instructor-led curriculums tend to provide instructionist classroom learning experiences. The elective curriculum itself provides students with a constructivist learning experience. But to succeed in exploring and developing their pathways, students need appropriate aids and resources to facilitate this type of learning process.

Most students rely on a few sources of information to forge their pathways. Academic advising and counseling has long offered students help with making academic and career decisions through conversations with trained experts. It is a valuable but scarce resource on most campuses. Not all students can have timely access to advisors. Students who might benefit the most are less likely to seek out advising staff, which raises equity concerns [31]. A more widely used source of information on campus is informal conversations with more senior students, but information gathered through this approach is often homogeneous and biased based on a student's social network [72]. Nevertheless, there may still be opportunities to leverage social learning from the academic choices of peers. Historical course enrollment data are usually only accessed by specialized staff and institutional leaders for high-level reports on academic progress, but recently, academic pathway data have been used to generate insights for instructors, students, and members of the research community [3, 55]. We are not aware of any application that provides students with personalized visual representations of archival course pathways, which could convey authentic, temporal information about course and major choices to inform and inspire current students.

We developed *Pathways*, a system to support academic decision-making at scale by visualizing campus-wide student pathway data to promote equal access to unbiased information about academic choices. The design of *Pathways* is grounded in psychological theories of interest: interest exploration is a fundamental component in constructivist learning [25]. It is also an important activity for students in the development of their individual and shared values [13, 33] and identities [60] during university education, which are influential throughout students' lives beyond finding a job and other utility-driven choices [24]. *Pathways* uses data storytelling techniques to provide information on course pathways through a series of interactions and visualizations that encourage students to explore. Data storytelling is commonly used in journalism to communicate insights from data with a specific goal [67, 68]. It can help convey complex information and guide exploration in a cognitively well-paced and compelling manner. While we investigate interest exploration among undergraduates, this approach extends to other contexts, including helping working learners navigate their learning journey outside of formal education.

This research makes three primary contributions: first, we develop an interactive and scalable tool that promotes interest exploration and supports academic decision making; second, we present empirical insights into challenges that impede interest exploration during academic decision making; and third, an empirical validation of the tool offers insights into how it supports students in exploring their interests and academic pathways.

2 BACKGROUND

2.1 Interest Exploration

Students who start their university education and realize how many options are available to them tend to ponder an important question: what am I interested in? Past research shows that a student's interest significantly contributes to their academic decisions [35] and career choices [46, 75]. Models of interest provide a theoretical foundation to study pathway exploration. The process of constructing pathways can serve as an affordance for students to reflect on their interests.

Past research on the psychological construct of interest distinguishes between situational and individual interest [36, 37]. Situational interest is a psychological state triggered by external cues. Individual interest is a trait-like preference that persists. The model of interest development posits that an individual's situational interest in a topic is first triggered by a particular situation [38, 62]. As they engage with the topic over time, their situational interest transitions into an individual interest that becomes a part of their personal characteristics. The model aligns with a constructivist understanding of interest, because it emphasizes the developmental process of interest instead of considering it to be fixed [25]. Accordingly, we define interest exploration in our research as the intentional effort that a learner puts forth during the initial stage of interest development when situational interest is triggered. Interest can also be triggered when one already has an individual interest [63]. The triggering process can facilitate specific aspects of an established interest to be deepened and evolved. Interest exploration is therefore not constrained to new college students with limited knowledge of what they want to study; it applies to students at all stages of their academic journey.

The design of *Pathways* accounts for the transition state between situational and individual interest for college students. Its goal is to help students not only find pathways based on individual interest, but also, explore situational interest by getting exposed to courses and pathways they would not have known about otherwise.

2.2 Digital Tools for Academic Decision Making

Every college and university provides an online course catalog system where students can look up course information. Students can browse courses offered by different departments and search for courses by course code, keywords in the course title, and instructor. The latest generation of course catalog systems, such as *Atlas* at the University of Michigan [53] and *Carta* at Stanford University [76], provide significantly more comprehensive information about courses that draw on data from official registrar records and student evaluations of teaching. These systems display information including grade distributions, official student reviews, estimated time spent on the course, and common courses taken before, with, or after a given course.

In recent years, historical course enrollment data has gained traction as a novel source of information to understand and support academic decision making [43, 74]. They can be used to reveal complex decision patterns that students take for sequential course navigation [3]. Moreover, applying computational methods to these data can suggest relations of courses through co-enrollment records [52, 57], which makes them powerful at course recommendation [55]. The availability of enrollment records at most institutions opens up new opportunities to support students' academic decision making at scale, especially for constructing pathways. For example, Shao and colleagues [70] developed a system that recommends courses across multiple semesters to students for degree planning. We build on prior work in this domain by developing a tool that makes historical course enrollment data accessible and interpretable to students so that they can explore potential pathways.

Most stand-alone, student-facing tools that support academic decision making follow a prescriptive advising model [41]. Crookston [22] describes it as an approach similar to how a doctor would treat a patient. Students receive "prescribed" answers based on practical needs, which tend to concern registration, record keeping, and generating recommendations. For example, one of the most extensive digital advising system can support students in choosing courses and majors, scheduling, and monitoring degree progress [58]. These systems, while important, only in part assist a student's personal pursuit. We opted for a developmental advising model in *Pathways*, which emphasizes helping students become aware of their changing self and realizing their potential [22]. Most academic decision support systems implicitly follow an assumption that interest is stable by treating it as an input for future choice prediction. In contrast, we recognize the changing state of a student's interest. We frame finding pathways as a chance for students to reflect on their changing interests. Moreover, we intentionally do not "prescribe" choices, but create interactive visualizations to encourage students to take on an active role in exploring options.

2.3 Data Storytelling

A common concern with data-driven tools that are exploratory in nature is that they are not actionable or interpretable enough to benefit students [21, 78]. Students may get lost during exploration and not be able to extract key insights from the visualizations.

Data storytelling is a powerful technique to communicate complicated information in data and convey ideas to its audience [8, 23, 44]. Specifically, it embeds narrative and structure in a series of data visualizations to highlight insights with an overarching goal [28, 68]. We adopt an "interactive slideshow" framework for data storytelling in *Pathways*, which walks students through steps like a slideshow [77]. The overall story-line is author-driven but with the overarching goal of promoting interest exploration. In each step, it takes a reader-driven approach for students to explore at the level of courses and entire pathways to ensure a high degree of flexibility while following the overarching narrative.

The concept of data storytelling is in popular use today to report news through mass media [17, 32] and it has been adapted to other areas (e.g., business [45, 49]), but few studies have applied it in educational contexts [29]. Chen and colleagues [20] designed narrative slideshows to help educators explore learning patterns in massive open online courses. Martinez-Maldonado and colleagues [50] organized data in layers of storytelling in a multi-modal learning analytics system to explain insights to teachers and students. These studies suggest that data storytelling can be successfully applied in educational settings to improve learning experiences, but more work is needed to develop theoretically grounded and learner-centered systems that demonstrate the effectiveness of data storytelling as a communication technique. We adopt data storytelling as an approach in *Pathways* to promote interest exploration through pathway-finding in a guided manner.

3 FORMATIVE STUDY

The process by which students consider courses for enrollment is complex and remains largely unobserved [19]. We therefore conducted a formative study to investigate the following two research questions: What challenges do students face exploring their interests during course consideration? (**RQ1**) How do students identify and develop their interests through courses? (**RQ2**) We will use our findings to draw design ideas that can support students to explore their interest and make academic decisions.

We conducted semi-structured interviews with twelve participants from a US research university (nine women and three men, nine Asian and three White students, average age 19.8, SD = 0.7). The interview protocol is available on OSF https://osf.io/g3fdr/. Participants were undergraduate students in various years (one first-year, two second-years, eight third-years, and one fourth-year student) and programs (two in Human Ecology, five in Agriculture and Life Sciences, four in Arts and Sciences, and one in Engineering). To analyze the interview recordings, we used deductive coding based on our two research questions. This led us to two main findings that informed our design decisions.

Finding for RQ1: Students face an overwhelming number of competing factors during course consideration. Participants consider many factors when deciding what courses to take, including interest (12 out of 12 participants), degree requirements (12/12), time schedules (10/12), class capacity (6/12), practical skills to be learned (6/12), instructional style (5/12), workload (5/12), difficulty (4/12), and grades (4/12). Participants felt "overwhelmed" by the many factors they need to consider during course selection. In order to cope, participants postponed consideration of certain factors to subsequent stages of their decision-making process, allowing them to focus on one or two factors at a time. However, the order in which participants considered factors varied. For example, P2 valued timing the most, so she excluded all early morning classes from the start. P8 only considered class difficulty when she made her final decision, whereas P3 considered difficulty from the beginning and only looked into easy classes. It shows that participants have different priorities that, if they dominate over their interest, can limit the scope for exploration. It is therefore important to preserve interest as a priority in students' course consideration process to encourage wider exploration.

Studies of multistage decision-making processes find that people expand their consideration set at an early stage, and then narrow it down in later stages [2, 6, 14]. If a student considers too many competing factors early on, they risk having a small consideration set to start with. One study found that students at one university considered a surprisingly narrow set of courses in their initial course consideration process [19]. Our interview findings suggest that factors such as difficulty can discourage students from pursuing their interests early on, which leads them to consider a narrower set of courses. We therefore intentionally designed *Pathways* as a tool that focuses on students' early decision stages with the goal of expanding students' awareness and consideration of course options to explore more before excluding them from consideration. The primary feature of *Pathways* thus begins with a topic-centered search to give students a space to freely articulate their interests. To prioritize student interest, we intentionally kept the tool simple, omitting features for course planning or course review to avoid cues that highlight other factors, including requirements, scheduling, difficulty, and instructional quality.

Finding for RQ2: Students identify their interests through serendipitous discoveries. Many participants developed a broad sense of their interests from daily interactions with family members (9/12), close friends (4/12), and public media (4/12). In college, taking courses enabled participants to systematically investigate their interests. As a result of taking more and more courses, most participants modified their initial interests or identified new ones (11/12). In retrospect, their interests often changed or developed unexpectedly—for many, a serendipitous experience. For example, P7 was initially interested in food science because "*it sounds cool*". But after taking several relevant courses, she found out that she was not as excited about biology experiments as her peers. It was then a course in food and consumers that helped her realize that she was actually interested in the marketing side of the food industry, an area she initially had limited knowledge of. In retrospect, P7 identified her new interest through a pivotal course that was broadly related to food. P2 shared a similar experience: she has always been interested in social work for families and children and initially focused on the psychological aspects of this area. But a course about children and the law introduced her to the relevant legal aspects, and she fell in love with it. Although her general interest did not change, her interest developed by exploring different courses about it. These cases highlight how an unexpected shift in a student's interest is often a result of pursuing an existing interest. This finding inspires us to design for serendipitous discoveries in *Pathways* by layering unexpected information on top of expected information.

Prior work on recommender systems developed algorithms that can increase the chance of serendipitous discoveries by adding novelty and diversity in recommendations [18]. This way the recommender system does not generate only relevant information and avoids users being trapped in "filter bubbles". In the context of course recommendation systems, Pardos and Jiang [56] developed an algorithm that returns courses within a student's indicated area of interest but possibly offered by another department that uses different disciplinary vernacular for the same topic in the course description. However, their recommendations do not go beyond a student's stated interest. To recognize potential transitions and developments in a student's interest, *Pathways* aims to generate search results that trigger new situational interest. We developed a serendipity-promoting algorithm that shows both relevant information and diverse-but-irrelevant information to students. However, this approach entails a trade-off: users may be dissatisfied because irrelevant information lowers the perceived accuracy of the search result. Thus, we divided the search results into separate steps, embedding them in the data storytelling framework so that students are gradually guided to more diversified content. We also created explanatory narratives in each step to convey the tool's design intentions to students.

Recommendation algorithms are one way to facilitate serendipitous discovery. Another approach is to use design strategies that can enhance user perceptions of serendipity [59]. In particular, recent studies have argued that serendipity should be conceptualized as an affordance in the design of tools and users should be granted agency to adopt "serendipity strategies" as opposed to serving serendipity "on the plate" [9, 48, 64]. They recommend adding interactions that give users autonomy to actively explore serendipity. We therefore added an intermediate calibration step in the data storytelling framework to let students calibrate the algorithm and inform the direction of diversity in the final result. *Pathways* positions students as actively seeking out diverse results, rather than being coerced to view content that may not match their interest.

4 *PATHWAYS*

The *Pathways* application is an interactive course exploration tool designed to promote interest exploration among college students. Building on our formative study, we designed the system as an exploratory search tool that presents courses and pathways to inspire serendipitous discoveries beyond students' initial interest. We adopted a data storytelling framework to organize the exploration process into three interactive steps, each with a specific sub-goal [65]. The goal of step one (Search) is to have students reflect on what their interests are and be open to all kinds of ideas. The goal of step two (Course Matches) is to support exploration in the student's domain of interest by showing a pool of courses that are independently relevant to their stated interest. The goal of step three (Pathway Matches) is to create serendipitous discoveries in interest exploration by showing pathway visualizations

Figure 1: *Pathways* User Interface.

that contain a combination of topics relevant to the student's stated interest and unexpected topics inherent in a given academic pathway. Each step is narrated in an informal style to communicate what is being presented to resemble an informal chat with a mentor. Transitions between steps follow a read-and-proceed interaction paradigm (i.e. students click a button and new content appears below). An overview of the user interface is shown in Figure 1 and the *Pathways* workflow can be found on OSF: https://osf.io/g3fdr/.

Step 1: Search. First, the student describes their general interests in an open-ended response field (Figure 1.a). To underscore the exploratory nature of the tool, we opted for an open-ended prompt ("Describe in a few sentences what you are interested in studying") instead of a series of closed-ended questions about students' departments or majors. While closed-ended responses can provide useful information for inferring interests, they may also limit the expression of ideas to subjects that students are already familiar with or have seen offered at their institution. Below the prompt, a worked example of a lengthy search response is shown to nudge students to write more. A new example response replaces the old one when a student clicks on "Load another example". Below the example response, a large input text box encourages students to reflect.

Step 2: Course Matches. Based on the student's stated interest, the most relevant courses are shown in hexagons and color-coded by the degree of relevance (Figure 1.b). The student can filter the courses by credit hours and subject. Course descriptions are shown when the student hovers over a course. If they click on the course, the tool will take them to the institution's official course catalog page with details such as the course description, units, and instructors. The student can save a course to their dashboard if it spurred their interest. All courses shown in this step are directly relevant to the student's stated interest, but the results are not constrained to one department. Thus, the results are more diverse than what the university provides in the course information system, which shows courses by filtering through department categories. This step is intended to be a stepping stone to serendipitous discoveries before showing the student results that extend from their stated interest.

Step 3: Pathway Matches. The student is presented a pool of courses that do not directly match their stated interest but may still be of interest to them (Figure 1.c). They are asked to select one or more courses that they find interesting. This is a calibration step for the algorithm to learn about the student's potential interests before pathways are shown. It is also intended to grant students agency to seek out potential interests. Finally, a full pathway visualization of a previous student is shown to the student (Figure 1.d). Courses are lined up in hexagons from top to bottom by semester in a chronological order, with "Year 1" to "Year 4" indicated on the left side. This type of temporal visualization depicts how course-taking patterns evolve over time and evokes a personal story for the anonymous student pathway beyond the mere list of courses they took [77]. The hexagonal shapes create a compact layout for the student to view all courses at first sight without the need to scroll much, which conforms with Mayer's temporal contiguity principle [51]. Courses are color-coded by their subjects to emphasize the diversity of subjects in a pathway. Underneath the visualization, the left/right arrow buttons let the student flip through five pathways. While these pathways contain courses that are relevant to the student's stated interest in the first step and the calibration step, they also contain courses with minimal overlap across pathways to achieve high diversity. The student can save entire pathways or courses of interest into their dashboard.

The Dashboard Page. The exploratory search feature is complemented by a dashboard page where students can browse and manage the list of courses and pathways they saved. Students can also initiate a new search from the dashboard page.

5 ALGORITHM DESIGN PROCESS

This section presents the development of the search algorithm that supports the functions described above. The algorithm operates on an archival, de-identified dataset of historical course enrollments and a dataset of course information that is periodically updated. A known design challenge of smart systems is that the technical feasibility of the algorithm depends on the UX design and dataset characteristics [27, 79]. Following best practices in HCI [10, 11], we conducted design-driven data exploration to guide the algorithm design and match it to our UX design. We also intentionally designed the algorithm to be simple enough to ensure high algorithm interpretability in each step of the results, as well as establish a baseline for future iterations with more complex, intelligent algorithms. We now describe the data we used, how results are generated in steps 2-3, and the design-driven data exploration process that led to the final algorithm design.

In the *Course Matches* step, the model performs single-item recommendation for course items, similar to models for web page searches. The search is performed based on a course information dataset, obtained via an API from the institution's public course catalogue with the latest information about all courses on campus. The dataset contains attributes such as course title, course description, instructors, and number of units. We use the course title and description as the course corpus for information retrieval, following these steps: First, the search query from step 1 is preprocessed by spellchecking, lowercasing, lemmatizing, and removing common English stop words as well as specific course-description related stop words such as"course", "teacher", and "instructor". Second, query expansion is performed to improve match quality: if the search query is under 25 words, it is expanded by adding related words from NLTK synsets. Third, the search is performed by computing TF-IDF vectors for unigrams and bigrams found in the query and in the course corpus. We show the top relevant courses to users based on cosine similarity.

In the *Pathway Matches* step, the model recommends a set of pathways composed of courses. The search is performed based on an archival dataset of historical course enrollments for 4,398 students between 2008 to 2019. The data includes all courses that students took each semester until they graduated. It includes nearly all course subjects that a student can take on campus (179 out of 186 subjects offered). Diversity in academic choice patterns can offer valuable insights to current students: instead of only learning about course choices from peers in their social network, students gain equitable access to information about academic decisions of diverse peers. A common challenge with data-driven applications like *Pathways* is that it can be unclear what outputs it will produce in the absence of user input during the initial design phase. We designed our

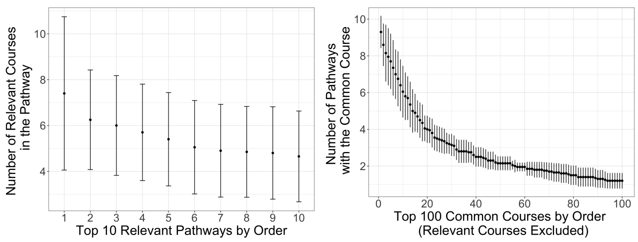

Figure 2: Left: Number of relevant courses (±*SD*) in the top 10 relevant pathways averaged across all search queries. Right: Number of pathways (±*SD*) in the top 100 common courses averaged across all search queries.

algorithm without knowing what students might write as search inputs. To address this challenge, we created twenty search queries to gain an initial sense of how the algorithm may work. We curated a diverse set of search queries in terms of their subject matter, length, colloquial style, and student motivations. Two example queries we used for testing are *AI for social good* and *I'm interested in classes that apply fine arts to something computational, like computational architecture or digital arts*.

To develop a serendipity-promoting algorithm, we created two metrics to measure both relevance and diversity in the results. We define pathway relevance by the number of relevant courses in a given pathway, where the top 50 courses with the highest cosine similarity are counted as relevant. We then define diversity across pathways by the number of common courses that a set of pathways contain, excluding all relevant courses. The fewer common courses there are, the more diverse the set of pathways is. We aim for an algorithm that achieves high pathway relevance and high diversity across pathways, and we conducted data exploration to understand our dataset in terms of these two metrics.

First, we investigated how many relevant courses a pathway contains for a given search query. We computed pathway relevance for our twenty example search queries and obtained the top ten relevant pathways in order. Figure 2.a shows that even the most relevant pathways contain only seven relevant courses on average. Students in our dataset take an average of 43 courses (SD = 8) throughout their undergraduate studies. Thus, at most 16% of courses in a student's pathway are directly relevant to a search query. This implies that even the most relevant set of pathways contains a high proportion of other courses for students to explore. Therefore, for a given search query, we design the algorithm to select the top ten pathways in terms of pathway relevance. This practically guarantees unexpected encounters with courses and topics while maintaining the highest degree of pathway relevance.

Every pathway with high relevance contains irrelevant courses, but it is unclear whether these courses will trigger student interest. We identified and adopted two approaches in our algorithm design to increase the chance that the presented pathway triggers student interest: One approach is to maximize diversity across pathways to broaden the student's exposure to various courses. The other approach is to prioritize the pathway that contains courses related to the student's potential interest. This raised the question: how many diverse courses are in pathway results for a given query? We computed the number of common courses excluding relevant courses in the top ten relevant pathways for all twenty example queries (Figure 2.b). The long tail in the distribution suggests that a large variety of courses are represented in even a small number (10) of pathways. Figure 2.b also shows that more than half of the top relevant pathways have approximately 15 courses in common, even after excluding relevant courses that many of these pathways contain. This indicates that common courses are an issue in pathways and may negatively affect the diversity of pathways. However, courses may also be common across pathways because they are good matches for a student's interest, suggesting to students that peers who share their interest mostly took the following set of courses. Therefore, in the calibration step, we ask students to pick courses of interest from a pool of top common courses among their relevant pathways. Based on the student's selection, the algorithm then selects a set of five pathways to achieve the highest diversity across pathways among those with the highest pathway relevance. Pathways are shown to the student in the order that prioritizes those with the most courses of interest selected by the student.

6 EVALUATION STUDY

We conducted an evaluation study to test the usability of *Pathways* and to understand academic exploration behaviors. Given that the goal of *Pathways* is to promote interest exploration during academic decision making, we pose the research question (**RQ1**): does *Pathways* support students to explore their interests? We also seek new insights into how interest exploration and academic decision making can be supported using *Pathways* as a research probe [40]. We observe how students react to *Pathways*, and especially to the design strategies we embedded in it, to answer our second research question (**RQ2**): how do students use *Pathways* to explore their interests and academic pathways? Answers to these questions will further our understanding of the needs and desires of students, and generate design ideas for future refinements of *Pathways* and related tools to support academic decision making.

6.1 Methods

6.1.1 Participants. We recruited fifteen participants from the university where the tool was developed: eight self-identified women and seven men; eight Asian, five White, one Hispanic, and one Asian-White mixed-race student; three first-year, seven second-year, three third-year, and two fourth-year or above students; three transfer students. Nine participants had already declared a major: four in information science, three in communications, one in applied economics and management, and one in computer science. We use a standard sample size for an initial, exploratory system evaluation in the HCI literature [16].

6.1.2 Procedure. We set up the evaluation study during the course pre-enrollment period, a time when students are highly engaged in academic decisions. The study was conducted one-on-one and online. First, participants were asked to explore *Pathways* for 10-20 minutes while using a think-aloud protocol [42]. Then, the experimenter sent over a short anonymous survey to solicit background information and a quantitative evaluation of *Pathways* to mitigate attitudinal biases that may be induced by the presence of a researcher. Due to connectivity issues, one participant was unable to fill out the survey. In the end, a semi-structured interview was conducted that lasted 20-30 minutes. Participants were asked about their general experience of *Pathways* and their personal interests and academic plans. The study sessions were video recorded, including the think-aloud exploration and semi-structured interview, but excluding the period when participants filled out the survey.

6.1.3 Survey Measures. Participants rated their agreement with seven statements about *Pathways* on a 7-point Likert scale (1 = Strongly Disagree, 2 = Disagree, 3 = Somewhat Disagree, 4 = Neither Agree nor Disagree, 5 = Somewhat Agree, 6 = Agree, 7 = Strongly Agree). The statements (shown in Figure 3) are about how well the serendipity-promoting algorithm shows relevant and diverse results (3 items), how the results translated into course discoveries (2), and about future use and recommendation intentions (2).

6.1.4 Interview Protocol and Analysis. In the interview (protocol in OSF: https://osf.io/g3fdr/), participants were asked to compare between *Pathways* and the institution's official course catalog system, *Roster*, that has standard features that most institutions provide today. Participants were also asked to describe what they would use *Pathways* for. The questions were designed to examine if *Pathways* is an improvement over the most commonly used system in colleges and if participants can discern how *Pathways* is different from a course catalog system. Participants' feedback was contextualized within their stated interests and situation so we can understand how they reached their evaluation of *Pathways* depending on their specific context. This allows us to investigate where *Pathways* was successful and where its limitations lie. We conducted a thematic analysis [12] using the collected recordings. The research team used line-by-line coding and generated memos individually. We analyzed the codes and memos together, looked for patterns across all interviews, and grouped them into themes. During this collaborative analysis, we discussed all key points in the memos for which there was disagreement or which were identified by one but not other researchers. We then reviewed the interview recording and discussed inconsistencies with the themes. Themes that did not contain enough codes for support were removed or combined into other themes. This was an iterative process that lasted until final agreement was reached.

6.2 Finding: Performance of *Pathways* (RQ1)

We present findings for RQ1 about how well *Pathways* supports students in exploring their interests. We organize findings into three parts: First, we evaluate if participants perceived the serendipitous-promoting algorithm as generating both relevant and diverse results. Second, we evaluate if the generated results led participants to discover courses that are both interesting and unexpected. Third, we evaluate if students would use the tool in the future and recommend it to their peers. We use evidence from the survey (Figure 3) and interviews in presenting our findings.

6.2.1 Relevance and Diversity in the Results. We examined participants' subjective evaluation of the results separately in the *Course Matches* step and the *Pathway Matches* step. For the first one, 13 out of 14 participants agreed (Median = 6, SD = 1.2) that the course matches were relevant. When using the tool, all participants saved at least one out of the ten courses to revisit later in their Dashboard.

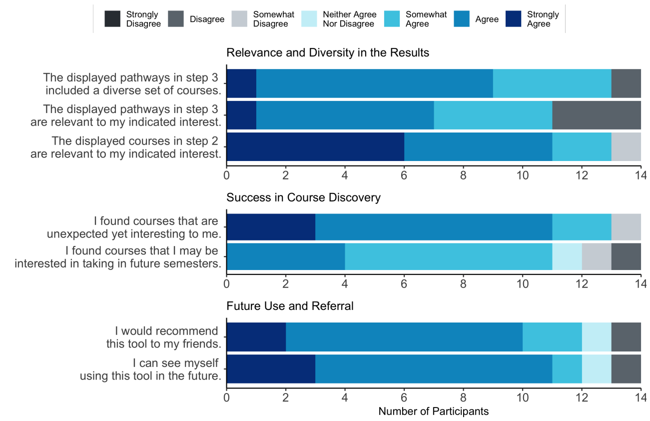

Figure 3: Survey responses in the evaluation study (n=14).

Most participants (10/14) reported that they saw courses related to their interests, including ones they took before, ones they planned to take, and notably, ones that they had not heard of before. This suggests that even a basic relevance match between a participant's search query and course descriptions can increase student awareness of course offerings and create unexpected encounters that may help develop their current interest.

For the *Pathway Matches* step, most participants still agreed that courses were relevant (11/14, Median = 5, SD = 1.7), though the level of agreement was slightly lower than in the *Course Matches* stage. This drop is expected as academic pathways include courses that are not closely related to a participant's indicated interest, which can affect the overall relevance of the result. In fact, 13 out of 14 participants agreed that pathways exposed them to a diverse set of courses (Median = 6, SD = 1.1). These findings highlight the competing goals of closely matching user needs with relevant information and extending beyond user needs with less relevant information. Thus, the inclusion of unexpected results needs to be well-balanced. Overall, the results suggest that *Pathways* presented students with results that balanced relevance and diversity as intended.

6.2.2 Success in Course Discovery. The goal of *Pathways* is to help students find academic interests through viewing relevant and diverse courses. We therefore examined if our approach enabled successful course discovery. Most participants agreed (13/14) that they found courses to take in future semesters (Median = 6, SD = 1.0). Moreover, they also agreed (13/14) that they found courses that are unexpected yet interesting to them (Median = 5, SD = 1.2). This finding is supported by evidence from the interviews. For instance, P12 remarked on how his situational interest was triggered in an unexpected way: "*I didn't even think about (this class) when I was initially saying what I was interested in in the first step but it popped up and I was like okay this is something I can see myself doing.*" Overall, the results suggest that *Pathways* was effective in stimulating student interest and creating serendipitous encounters during academic exploration.

Participants also provided feedback on the interaction design and visualizations to discover courses of interest easier. In the calibration step of *Pathway Matches*, five participants commented that they were not interested in any courses presented to them and thus cannot calibrate the algorithm. They therefore suggested adding a "load more" feature to browse through more courses for the calibration step. In the pathway visualizations, three participants raised the concern that color-coding courses by subject makes the pathway too "*noisy*" (P1) or "*distracting*" (P3) because there are too many subjects (Mean = 15 in a pathway, SD = 3.1). They therefore suggested highlighting information in pathways based on broader categories, such as highlighting STEM courses (P13) or courses often taken in the final college years (P6). Additionally, participants appreciated the compactness and sense of liveliness that displaying courses in hexagons evokes, but the constrained width of the hexagon makes course names hard to read and courses of interest less noticeable. P10 suggested that seeing courses as a list may be better. This may be particularly true for the visualization in *Course Matches*, because arranging search results by relevance in a list format is a common design pattern for search engines. Overall, the interviews provided clear feedback for refinements to the visualization for future iterations of *Pathways* to improve its usability.

6.2.3 Future Use and Referral. Most participants (12/14) agreed that they would like to use *Pathways* in the future (Median = 6, SD = 1.3) and would recommend it to friends (Median = 6, SD = 1.3). The interview data offers further insight into their reasons and potential uses. First, participants found it to be a useful, complementary tool to the institution's official course catalog system, *Roster*. Participants described *Pathways* as a way to help them find potential courses of interest, while *Roster* would be used subsequently once they already know what courses to take. P5 explains the difference: "*(With) Roster you had to do the research. This does the research for you.*" Second, participants saw *Pathways* as a tool to improve information equity and mitigate biases, because it allows them to get course-taking advice outside of their social circle: "*I'm actually interested in data science, but like I didn't find any people in the data science field to ask them like which courses are best to learn data analysis.*" (P6). Participants appreciated the ability to gather information that is impartial compared to their friends' point of view: "*If I ask someone else maybe they're more into designs they may recommend design classes but I'm gonna do like front-end and back-end engineering stuff so it's not really that relevant.*" (P13).

Participants came to appreciate connections between courses in academic pathways and engage in longer-term thinking about their academic plans with *Pathways*: "*You can see what courses are used to like built based on higher-level courses. That's definitely something I haven't seen anywhere else.*" (P13). The forward-thinking design of *Pathways* is especially useful to first-year students who may know little about what a four-year degree program entails: "*I think something... you don't realize is you kind of have to have a four-year plan when trying to pick classes, but when you come as a freshman like 'What I think of my four years' is definitely a lot.*" (P8). Multi-semester planning is also important to transfer students like P11. She commented that thinking strategically about more than one semester at a time is more common for transfer students, because they have more courses to cover in fewer semesters once they arrive at a new institution.

6.3 Finding: Reactions to *Pathways* (RQ2)

This section presents findings related to RQ2 about how students explore their interest and find courses using *Pathways*. We focus on the qualitative evidence from the evaluation study and organize findings around three claims: (1) Students struggle to articulate their interest; (2) Students can be resistant to interest exploration; and (3) Students gain new insights from pathway visualizations.

6.3.1 Students struggle to articulate their interest. Students are asked to indicate their interests in the search prompt on *Pathways*, but when participants faced the large open-ended text box, they expressed a range of attitudes toward it. Half of the participants (7/15) thought the open-endedness of the text box raises uncertainty over how to describe their interest. They also felt reluctant to articulate in full sentences. They would have preferred the search to ask more specific questions that can be answered in short phrases, such as "what major are you interested in" or "what was your favorite class". However, other participants appreciated the large space to be "*wordy*" (P11) about their interests. Participants who overlooked the example queries tended to spend less time thinking about what to write in the search box and typed in keywords such as their majors. Yet participants who closely read through the full-sentence example queries found them extremely helpful as a guide for what to write about and wrote in full sentences. They generally spent more time thinking and wrote more about their interests. This aligns with the design intention of *Pathways* to promote deeper reflection through articulating interests, even if it makes students feel uneasy. What came as a surprise is how much participants relied on example queries not only to determine how much to write in their query but also what to search for. Several participants (5/15) found topics in the example query interesting and decided to search for those or add those to their own interest. This may indicate how students find it taxing to reflect on their interests and rely closely on what is presented as long as it generally aligns with their interest. This also suggests that examples in the user interface can potentially be used as light-touch interventions to influence students' situational interest.

Many participants (9/15) revised their search query after viewing the search results. Among them, four participants realized that the interest that they indicated was too general based on the course results, so they wrote longer sentences to explain their interest. For example, P13 initially only wrote "*security*" in the search prompt, but then added details: "*I'm interested in system security, something in the computational field, and mechanical physics*". This shows how *Pathways* uses course search results to nudge students to reflect on and concretely articulate their interests. Five participants were inspired by search results that were not directly relevant to what they indicated. P10 initially wrote "*I'm interested in classes related to computer science, especially machine learning and human computer interaction. I'm also interested in data science.*", but after seeing a particular pathway, she went back and changed her search query into "*I'm looking for advanced computer science classes relating to computer architecture and robotics.*" This shows how *Pathways* can trigger new situational interest by including seemingly irrelevant information to students. Both cases show how courses and pathways can serve as "tangible objects" that allow students to "play with" their interests and develop a new understanding of it.

6.3.2 *Students can be resistant to interest exploration.*

All participants liked the *Course Matches* feature because it directly connected to their search query, but their attitudes were mixed about the *Pathway Matches* feature. Many participants (10/15) appreciated having *Course Matches* to set the stage for the more complicated *Pathway Matches* to make them "*easy to understand*" (P3). Some (3/15) explored the calibration feature to show different *Pathway Matches*, commenting that it makes pathway visualizations "*versatile*" (P4). However, we also saw that some participants (6/15) were not as enthusiastic and showed confusion or resistance when they were prompted by the calibration feature to explore or when they saw the diverse pathways. We found two reasons for this resistance.

First, some participants (3/15) are open to interest exploration, but only within the bounds of their requirements. They expressed difficulty reconciling the need to satisfy requirements with exploring their interest outside of what is required, describing it as a "*waste of credits*" (P5). For example, P6 tried to recognize which academic major a student pursued based on their pathways, such as "*this person is heavy business*" or "*this is an information science and economics person*"; once identified, P6 would move on to the next pathway until she found one that matched her requirements. While other participants did not consider fulfilling requirements their main priority, many of them (9/15) indicated a desire to see the major of the student whose pathway they were viewing, so that they could determine if the courses they found interesting would also satisfy their own degree requirements. They also expressed a desire for *Pathways* to let them specify their major in the search prompt so that the results could filter out pathways with different requirements. While this suggestion may serve some students' needs, it conflicts with *Pathways*' goal to facilitate broad exploration. P12 commented on the question of whether student major should be reported and incorporated in generating pathway results: "*There are pros and cons. They will filter in more classes in your major but the same time it can be very limiting. There can also be so many pathways that you may not even know you're interested in before and maybe over time you may grow interest in.*"

The other reason that some participants (3/15) chose not to explore is because they believe they had already committed to an interest. P10 expressed that, "*seeing all classes together (in a pathway) takes the focus away from what I'm interested in.*" Because most students formalize their interest in terms of majors and minors, participants behaved differently depending on whether they had already determined a major or minor. For example, P14 is a second-year student who is searching for a minor. He is interested in applying computer science in other disciplines, possibly in linguistics, philosophy, business and education. P14 reported being fascinated by the tool and he spent a long time exploring various outcomes. Participants like P14 mainly used *Pathways* to better understand the academic pathway of prospective majors or minors that they are interested in. However, P12 was less excited about the tool. Even though he is a freshman, he was already determined to declare information science as his major: "*It's a little different for me (than other freshmen) because I kind of already have an idea of what I want to do. [...] I feel like for students who are more undeclared... it may be more useful*". However, what these participants are often not aware of is that exploration is an ongoing process, even with an established interest. For example, P9 is a fourth-year student majoring in communication and is very open-minded about exploration. She described her motivation for exploration as follows: "*I think this application would be a really cool way to connect to what I'm pursuing for next semester. [...] Like maybe I don't want to just do marketing communications. Maybe I want a photography class. [...] Maybe there is a class that incorporates photography (with communication) but I don't even know and so I think this app being able to connect would be super cool for exploration.*" This case highlights that those who generally enjoy exploration tend to be open to topics that can provide new perspectives on their existing interest even if their interest is well-established.

6.3.3 *Students gain new insights from pathway visualizations.*

When participants browsed through visualizations of different pathways, we were surprised to find that they would extend our working definition of diversity. While we defined diversity only in terms of the variety in course topics across pathways, participants pointed out diversity in terms of when courses were taken, not just what kinds of courses were taken. They noticed many course-taking patterns in the pathways that surprised them because the patterns deviated from their understanding of what a pathway should look like. Specifically, they noted the following unexpected patterns: some students spent the first year primarily taking courses in one field but then transferred to a completely different field in their second year (noticed by P9, P15); some students took an introductory course in an a secondary field in their final year (noticed by P8); some students took courses in topics that look completely irrelevant to their field (noticed by P6, P11, P12, P15); and some students only took three courses in a semester when the usual number is around five (noticed by P11). Many participants expressed surprise and even asked if these pathways were from actual students. Seeing these course-taking patterns broadened participants' understanding to the greater possibilities and options for a pathway.

In addition to noticing different patterns, participants also noticed similarities across pathways. In most pathways, students took courses from a diverse pool of subjects in their early academic years, and later narrowed them down to one or two primary fields. This sends an important message to students: they can take diverse courses early on to explore what works for them, or as P9 put it: "*What this really shows is that all freshmen kind of have a very all-over-the-place schedule cause they're trying to figure things out. No one knows right off the bat that Oh this kid is gonna do CS, this kid is gonna do business. [...] I think a lot of students don't understand is that [...] you should take time to figure things out. Seeing this really motivates you. [...] You can really discover anything*." The patterns that participants identified in pathway visualizations are not intentionally highlighted in our original design, but they reveal insights that are at the core of what *Pathways* tries to achieve.

7 DISCUSSION AND CONCLUSION

We developed *Pathways* to support academic decision making at scale. We adopted a developmental advising approach to encourage students to explore their interests by visualizing past students' course enrollment records, using a data storytelling framework, and a serendipity-promoting algorithm. Our evaluation study shows that *Pathways* can help students discover new courses and potential pathways of interest. We found that many students struggle to articulate their interest, and that presenting courses in the form of pathways can help students formulate and express their interests. We also found that students tend to explore their interests within the bounds of degree requirements, or they view exploration as unnecessary once they declare an academic major. When students explore pathway visualizations, they realize that they can endeavor diverse pathways themselves. The study's findings raise broader theoretical and applied implications concerning academic interest exploration and data communication.

We conceptualize academic decision making as a constructivist learning process that is enabled by the elective system in U.S. higher education. Students need to create academic pathways that not only satisfy degree requirements, but also align with their developing interests and goals. To support this constructivist learning process, we did not design *Pathways* to be an AI "expert" system for academic decision making. Instead, it facilitates learning from peers using minimal AI to select peers who address students' needs. We found students actively reflecting on and evolving their interest in the process of using the tool. Students "observed" the academic decisions of others, represented as pathways, and compared them with their own academic decisions. This demonstrates how *Pathways* supports constructivist learning for academic decisions, and motivates further research into advising tools that support this kind of learning process with a high degree of student agency. Learning to make these decisions is a useful skill for life-long learning and career choices beyond college [66].

We now consider limitations of this research and opportunities for addressing them in future work. First, the historical course enrollment data we used was a sample of undergraduate students at the university that did not include all academic majors. This could have introduced biases, limited the diversity of results, and impeded serendipitous discoveries. While we did not notice problems along these lines, we plan on expanding the data set in future studies. Second, the evaluation study we conducted was in a controlled context rather than a field deployment, which limits the external validity of our findings. Participants used *Pathways* during the pre-enrollment period, but they did not use it "in the wild", which may influence their attitudes and behaviors. Third, the sample of participants was recruited from a channel that reaches a relatively homogeneous sample of students with interests in information science and communication. Their evaluation of the tool and its algorithm may be shaped by their interests. While our sample size is standard for an exploratory study with an initial design [16], the complexity of pathway results might require a larger, more diverse sample to gain generalizable insights about the performance of *Pathways*, and what it tells us about student interest exploration.

Our study reveals unexpected tensions in interest exploration during academic decision making. We were surprised to see how challenging it is for students to articulate their interest in ways other than stating their major. Majors are a useful heuristic for students to think about their interest and it helps students stay focused and graduate on time. The need to pick a major motivates students to explore if their interest fits [71]. Yet majors also constrain interests. We found that students articulate interests in a more nuanced, flexible, and creative way than majors after browsing diverse courses and pathways. Courses and pathways offer students a new channel to envision and develop their interest beyond majors. For today's students, it is especially important to consider interests beyond institutional definitions of academic disciplines (majors/minors), which are broadly defined and less fluid compared to rapidly changing industry demands [7]. How to design prompts and scaffolds to help students learn to articulate their interest warrants further study. Moreover, considering interests beyond one's major can promote exploration after major declaration. Based on prior work on interest development, it is equally important to trigger new situational interest in later interest development stages as it is in early stages [34, 61]. The time of major declaration may separate different stages of interest development: once declared, students have an established interest. We found that only a few students with a declared major continued exploration using *Pathways*, but new situational interest was triggered when they found cross-disciplinary courses that connect back to their established interest. We recommend future research to investigate different types of triggers for different stages of interest development to support sustained exploration.

Finally, our study shows the value of making historical course enrollment data transparent to students. While most advising tools focus on developing prediction models to inform academic decisions [43], *Pathways* demonstrates that descriptive information about other students' choices can be equally, if not more, constructive to support decision making. This approach is extremely scalable because it requires only simple technology and course enrollment data that is available at virtually every university. These data hold valuable information about academic choices and what they entail. But how to strategically surface this kind of information to students in a meaningful way requires further research. In *Pathways*, we used data storytelling and visualization techniques to let students view and compare choice options across peers. We found that students identify patterns that are instructive to them. They notice that many students do not have a clear academic path in the first terms, and are surprised to discover that academic pathways are much more diverse than they presumed. While we cannot tell why students took certain courses at certain times, the normative message conveyed by these observations is compelling: there is no standard way of creating a pathway and it is normal to take time to figure it out. Prior work suggests that social norms emerge when individuals interact with others [69] and they get reinforced when interacting in a relatively homogeneous group [54]. Our work shows that when students interact with a wider group of students, even by viewing their pathways asynchronously online, they are encouraged to update their normative beliefs and explore options that could better suit their needs. Tools like *Pathways* provide channels to convey academic norms with authentic data. We call for more research on students' exploration and decision making to build a science of academic pathways.

ACKNOWLEDGMENTS

We thank the Cornell Offices of the Vice Provost for Academic Innovation, University Registrar, and Information Technology for supporting this work.

REFERENCES

[1] Jaison R Abel and Richard Deitz. 2015. Agglomeration and job matching among college graduates. *Regional Science and Urban Economics* 51 (2015), 14–24.

[2] Naeim H Abougomaah, John L Schlacter, and William Gaidis. 1987. Elimination and choice phases in evoked set formation. *Journal of Consumer Marketing* (1987).

[3] Geoffrey Angus, Richard Diehl Martinez, Mitchell L Stevens, and Andreas Paepcke. 2019. Via: Illuminating academic pathways at scale. In *Proceedings of the Sixth (2019) ACM Conference on Learning@ Scale*. 1–10.

[4] Elizabeth A Armstrong and Laura T Hamilton. 2013. *Paying for the Party*. Harvard University Press.

[5] Thomas R Bailey. 2015. *Redesigning America's community colleges: A clearer path to student success*. Harvard University Press.

[6] Rachel Baker and Gabe Avakian Orona. 2020. Gender and Racial Differences in Awareness and Consideration of Curricular Programs: Exploring a Multistage Model of Major Choice. *AERA Open* 6, 3 (2020), 2332858420937023.

[7] Ashok Bardhan, Daniel L Hicks, and Dwight Jaffee. 2013. How responsive is higher education? The linkages between higher education and the labour market. *Applied Economics* 45, 10 (2013), 1239–1256.

[8] Skylar Bayer and Annaliese Hettinger. 2019. Storytelling: A Natural Tool to Weave the Threads of Science and Community Together. *Bulletin of the Ecological Society of America* 100, 2 (2019), 1–6. https://www.jstor.org/stable/26607530

[9] Lennart Björneborn. 2017. Three key affordances for serendipity: Toward a framework connecting environmental and personal factors in serendipitous encounters. *Journal of Documentation* (2017).

[10] Kirsten Boehner and Carl DiSalvo. 2016. Data, design and civics: An exploratory study of civic tech. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2970–2981.

[11] Sander Bogers, Joep Frens, Janne Van Kollenburg, Eva Deckers, and Caroline Hummels. 2016. Connected baby bottle: A design case study towards a framework for data-enabled design. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. 301–311.

[12] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[13] Jere Brophy. 1999. Toward a model of the value aspects of motivation in education: Developing appreciation for particular learning domains and activities. *Educational psychologist* 34, 2 (1999), 75–85.

[14] Elizabeth Bruch and Fred Feinberg. 2017. Decision-making processes in social contexts. *Annual review of sociology* 43 (2017), 207–227.

[15] Erik Brynjolfsson and Andrew McAfee. 2014. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.

[16] Kelly Caine. 2016. Local standards for sample size at CHI. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 981–992.

[17] Andreu Casero-Ripollés, Silvia Marcos-García, and Laura Alonso-Muñoz. 2020. New Formats for Local Journalism in the Era of Social Media and Big Data: From Transmedia to Storytelling. In *Journalistic Metamorphosis*. Springer, 69–83.

[18] Pablo Castells, Neil J Hurley, and Saul Vargas. 2015. Novelty and diversity in recommender systems. In *Recommender systems handbook*. Springer, 881–918.

[19] Sorathan Chaturapruek, Tobias Dalberg, Marissa E Thompson, Sonia Giebel, Monique H Harrison, Ramesh Johari, Mitchell L Stevens, and Rene F Kizilcec. 2021. Studying Undergraduate Course Consideration at Scale. *AERA Open* 7 (2021), 2332858421991148.

[20] Qing Chen, Zhen Li, Ting-Chuen Pong, and Huamin Qu. 2019. Designing Narrative Slideshows for Learning Analytics. In *2019 IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, 237–246.

[21] Rianne Conijn, Luuk Van Waes, and Menno van Zaanen. 2020. Human-centered design of a dashboard on students' revisions during writing. In *European Conference on Technology Enhanced Learning*. Springer, 30–44.

[22] Burns B Crookston. 1972. A developmental view of academic advising as teaching. *Journal of College Student Personnel* (1972).

[23] M. F. Dahlstrom. 2014. Using narratives and storytelling to communicate science with nonexpert audiences. *Proceedings of the National Academy of Sciences* 111 (Sep 2014), 13614–13620. https://doi.org/10.1073/pnas.1320645111

[24] William Deresiewicz. 2015. *Excellent sheep: The miseducation of the American elite and the way to a meaningful life*. Simon and Schuster.

[25] John Dewey. 1913. *Interest and effort in education*. Houghton Mifflin.

[26] John Dewey. 1986. Experience and education. In *The educational forum*, Vol. 50. Taylor & Francis, 241–252.

[27] Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX design innovation: Challenges for working with machine learning as a design material. In *Proceedings of the 2017 chi conference on human factors in computing systems*. 278–288.

[28] Brent Dykes. 2016. Data storytelling: The essential data science skill everyone needs. *Forbes* 3 (2016).

[29] Vanessa Echeverria, Roberto Martinez-Maldonado, Roger Granda, Katherine Chiluiza, Cristina Conati, and Simon Buckingham Shum. 2018. Driving data storytelling from learning design. In *Proceedings of the 8th international conference on learning analytics and knowledge*. 131–140.

[30] David Epstein. 2021. *Range: Why generalists triumph in a specialized world*. Penguin.

[31] Kevin Fosnacht, Alexander C McCormick, Jennifer N Nailos, and Amy K Ribera. 2017. Frequency of first-year student interactions with advisors. *NACADA Journal* 37, 1 (2017), 74–86.

[32] Jonathan Gray, Lucy Chambers, and Liliana Bounegru. 2012. *The data journalism handbook: How journalists can use data to improve the news*. " O'Reilly Media, Inc.".

[33] Judith M Harackiewicz and Chris S Hulleman. 2010. The importance of interest: The role of achievement goals and task values in promoting the development of interest. *Social and Personality Psychology Compass* 4, 1 (2010), 42–52.

[34] Judith M Harackiewicz, Jessi L Smith, and Stacy J Priniski. 2016. Interest matters: The importance of promoting interest in education. *Policy insights from the behavioral and brain sciences* 3, 2 (2016), 220–227.

[35] Ellen Karoline Henriksen, Justin Dillon, and Jim Ryder. 2015. *Understanding student participation and choice in science and technology education*. Springer.

[36] Suzanne Hidi and William Baird. 1988. Strategies for increasing text-based interest and students' recall of expository texts. *Reading research quarterly* (1988), 465–483.

[37] Suzanne Hidi and Judith M Harackiewicz. 2000. Motivating the academically unmotivated: A critical issue for the 21st century. *Review of educational research* 70, 2 (2000), 151–179.

[38] Suzanne Hidi and K Ann Renninger. 2006. The four-phase model of interest development. *Educational psychologist* 41, 2 (2006), 111–127.

[39] Andreas Hirschi. 2010. Individual predictors of adolescents' vocational interest stabilities. *International Journal for Educational and Vocational Guidance* 10, 1 (2010), 5–19. https://doi.org/10.1007/s10775-009-9171-2

[40] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. 2003. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 17–24.

[41] Omiros Iatrellis, Achilles Kameas, and Panos Fitsilis. 2017. Academic advising systems: A systematic literature review of empirical evidence. *Education Sciences* 7, 4 (2017), 90.

[42] Monique WM Jaspers, Thiemo Steen, Cor Van Den Bos, and Maud Geenen. 2004. The think aloud method: a guide to user interface design. *International journal of medical informatics* 73, 11-12 (2004), 781–795.

[43] Weijie Jiang and Zachary A Pardos. 2020. Evaluating Sources of Course Information and Models of Representation on a Variety of Institutional Prediction Tasks. *International Educational Data Mining Society* (2020).

[44] Aisling Kelliher and Malcolm Slaney. 2012. Tell Me a Story. *IEEE MultiMedia* 19 (01 2012), 4. https://doi.org/10.1109/MMUL.2012.13

[45] Cole Nussbaumer Knaflic. 2015. *Storytelling with data: A data visualization guide for business professionals*. John Wiley & Sons.

[46] Frederic Kuder. 1977. Activity Interests and Occupational. *Choice-Chicago: Science Research Associates* (1977), 144–146.

[47] Robert W Lent. 2018. Future of work in the digital world: Preparing for instability and opportunity. *The Career Development Quarterly* 66, 3 (2018), 205–219.

[48] Stephann Makri, Ann Blandford, Mel Woods, Sarah Sharples, and Deborah Maxwell. 2014. "Making my own luck": Serendipity strategies and how to support them in digital information environments. *Journal of the Association for Information Science and Technology* 65, 11 (2014), 2179–2194.

[49] Vanessa Maritz. 2012. *Strategic use of data visualisation and storytelling in marketing research firms*. Ph.D. Dissertation. University of Johannesburg.

[50] Roberto Martinez-Maldonado, Vanessa Echeverria, Gloria Fernandez Nieto, and Simon Buckingham Shum. 2020. From data to insights: A layered storytelling approach for multimodal learning analytics. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–15.

[51] Richard E Mayer. 2002. Multimedia learning. In *Psychology of learning and motivation*. Vol. 41. Elsevier, 85–139.

[52] Gonzalo Méndez, Xavier Ochoa, and Katherine Chiluiza. 2014. Techniques for data-driven curriculum analysis. In *Proceedings of the fourth international conference on learning analytics and knowledge*. 148–157.

[53] University of Michigan. 2021. Atlas. https://atlas.ai.umich.edu/

[54] Karl-Dieter Opp. 2002. When do norms emerge by human design and when by the unintended consequences of human action? The example of the no-smoking norm. *Rationality and Society* 14, 2 (2002), 131–158.

[55] Zachary A Pardos, Zihao Fan, and Weijie Jiang. 2019. Connectionist recommendation in the wild: on the utility and scrutability of neural networks for personalized course guidance. *User modeling and user-adapted interaction* 29, 2 (2019), 487–525.

[56] Zachary A Pardos and Weijie Jiang. 2020. Designing for serendipity in a university course recommendation system. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. 350–359.

[57] Zachary A Pardos and Andrew Joo Hun Nam. 2020. A university map of course knowledge. *PloS one* 15, 9 (2020), e0233207.

[58] Elizabeth D Phillips. 2013. Improving advising using technology and data analytics. *Change: The Magazine of Higher Learning* 45, 1 (2013), 48–55.

[59] Tammera M Race and Stephann Makri. 2016. *Accidental information discovery: cultivating serendipity in the digital age*. Elsevier.

[60] K Ann Renninger. 2009. Interest and identity development in instruction: An inductive model. *Educational psychologist* 44, 2 (2009), 105–118.

[61] K Ann Renninger and Suzanne Hidi. 2002. Student interest and achievement: Developmental issues raised by a case study. In *Development of achievement motivation*. Elsevier, 173–195.

[62] K Ann Renninger, Suzanne Hidi, Andreas Krapp, and Ann Renninger. 2014. *The role of interest in learning and development*. Psychology Press.

[63] K Ann Renninger and Suzanne E Hidi. 2019. Interest development as a dynamic process in the workplace. *Vocational interests in the workplace: rethinking behavior at work* (2019), 39–58.

[64] Urbano Reviglio. 2019. Serendipity as an emerging design principle of the infosphere: challenges and opportunities. *Ethics and Information Technology* 21, 2 (2019), 151–166.

[65] María Teresa Rodríguez, Sérgio Nunes, and Tiago Devezas. 2015. Telling Stories with Data Visualization. In *Proceedings of the 2015 Workshop on Narrative & Hypertext* (Guzelyurt, Northern Cyprus) *(NHT '15)*. Association for Computing Machinery, New York, NY, USA, 7–11. https://doi.org/10.1145/2804565.2804567

[66] Josipa Roksa and Tania Levey. 2010. What can you do with that degree? College major and occupational status of college graduates over time. *Social forces* 89, 2 (2010), 389–415.

[67] Lindy Ryan. 2016. *The visual imperative: Creating a visual culture of data discovery*. Morgan Kaufmann.

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

[69] Sandip Sen and Stéphane Airiau. 2007. Emergence of norms through social learning.. In *IJCAI*, Vol. 1507. 1512.

[70] Erzhuo Shao, Shiyuan Guo, and Zachary A Pardos. 2021. Degree Planning with PLAN-BERT : Multi-Semester Recommendation Using Future Courses of Interest. 35, 17 (2021), 14920–14929.

[71] John C Smart, Kenneth A Feldman, and Corinna A Ethington. 2000. *Academic disciplines: Holland's theory and the study of college students and faculty*. Vanderbilt University Press.

[72] Tobias H Stark. 2015. Understanding the selection bias: Social network processes and the effect of prejudice on the avoidance of outgroup friends. *Social Psychology Quarterly* 78, 2 (2015), 127–150.

[73] Leslie P Steffe and Jerry Edward Gale. 1995. *Constructivism in education*. Psychology Press.

[74] Mitchell Stevens, Monique Harrison, Marissa Thompson, Arik Lifschitz, and Sorathan Chaturapruek. 2018. Choices, Identities, Paths: Understanding College Students' Academic Decisions. *Available at SSRN 3162429* (2018).

[75] Edward K Strong Jr. 1943. Vocational interests of men and women. (1943).

[76] Stanford University. [n.d.]. Carta. http://carta-beta.stanford.edu/

[77] Wibke Weber. 2020. Exploring narrativity in data visualization in journalism. *Data visualization in society* (2020), 295–311.

[78] Alyssa Friend Wise and Yeonji Jung. 2019. Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics* 6, 2 (2019), 53–69.

[79] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Reexamining whether, why, and how human-ai interaction is uniquely difficult to design. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–13.