Visualizing Multi-Criteria Evaluations of Application Data in University Admissions: Supporting Holistic and Collaborative Decision-Making

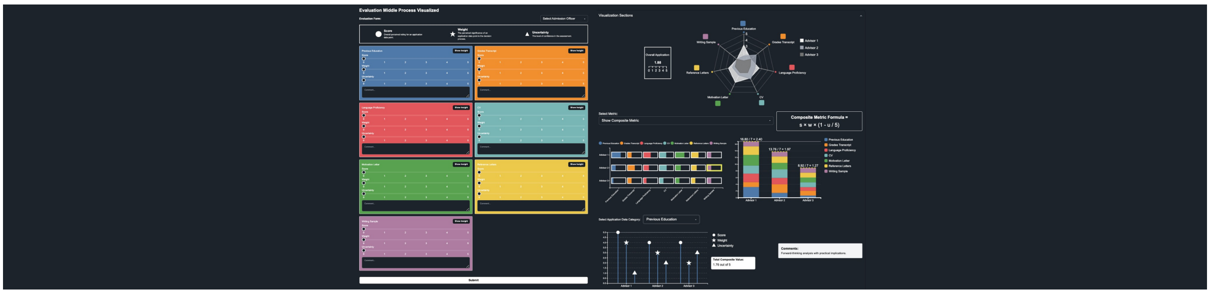
MASTER THESIS HUMAN-COMPUTER INTERACTION

Student Name: Samin Asnaashari (6388205)

Email: s.asnaashari@students.uu.nl

First Supervisor:
Name: Evanthia Dimara
Email: e.dimara@uu.nl

Second Supervisor:
Name: Christof van Nimwegen
Email: c.vannimwegen@uu.nl



Department of Information Technology

October, 2024

Abstract

University admissions processes are complex, requiring evaluators to assess multiple diverse data points—such as academic records, CVs, and motivation letters—to form a holistic view of an applicant's potential. However, the manual nature of these assessments can lead to inconsistencies, particularly when multiple evaluators are involved. This thesis presents **EvaluationViz**, a visualization tool designed to support multi-criteria decision-making (MCDM) in university admissions. The tool enables evaluators to assess applicants across seven key data points—**previous education, grades transcript, language proficiency, CV, motivation letter, reference letters, and writing sample**—using three key metrics: **Score**, **Weight**, and **Uncertainty**. EvaluationViz integrates interactive visualizations, such as **radar/spider charts, tabular panel bar charts, stacked bar charts, and lollipop charts**, to present application evaluation data in a clear, intuitive manner, facilitating more transparent, consistent, and collaborative decision-making. The tool's development was guided by a literature review on information visualization and MCDM, and further refined through a pilot study involving 9 admissions experts at Utrecht University, followed by a feedback session with 7 participants. The evaluation involved a controlled experiment with 12 participants, grouped into four teams of three. Each participant evaluated applicants in two phases: first individually, and then as part of a group. Each phase used different anonymized applicant profiles, but all participants applied the same evaluation rubric and decision-making task in both phases. In the first phase, participants assessed applicants without the use of visualizations, relying solely on provided forms. In the second phase, they repeated the assessment using EvaluationViz. Results revealed that participants struggled to synthesize the various application data points without visual aids, often focusing on isolated factors like grades and overlooking the integration of metrics. By contrast, in the second phase, participants reported that visualizations, particularly the radar chart, tabular panel bar chart, and lollipop chart, helped them better integrate scores, weights, uncertainties, and identity where in data these metrics affected the overall evaluation. One participant noted that the visualizations allowed them to "see where we placed the most weight and uncertainty in the decision," leading to more structured and evidence-based group discussions. The tool also reduced fragmented discussions, allowing participants to focus on the holistic assessment of each applicant. Despite these promising results, the evaluation was conducted using simulated data in a controlled environment, which limits its direct applicability to real-world fast-paced admissions processes. Future research should focus on testing the tool with actual applicant data, automating specific evaluation tasks, adding more dimension, integrating historical data, and enabling comparisons with similar applicants through visual encodings such as Sankey diagrams. By addressing these limitations, EvaluationViz has the potential to become a valuable asset in university admissions, fostering more efficient, data-driven, and equitable decisions.

Contents

1 Introduction

According to the *Oxford English Dictionary*, **"admission"** is *the process or fact of entering or being allowed to enter a place or organization* [51]. This concept can be further understood as a systematic procedure for evaluating applicants across different contexts, ensuring that they meet predefined criteria established within a specific policy framework. Through this framework, evaluators make informed decisions that result in the final judgment of entry or non-entry [34].

In educational settings, this process can be described as whether applicants meet the necessary prerequisite qualifications, have the potential to complete the program successfully, and are in line with the institution's goals. Typically, evaluators consider academic records, rank and credibility of the applicant's previous education, and motivation to assess the applicant's fit for the program [71].

A similar approach is also applied in professional settings, while here, the focus is shifted to evaluating resumes, required qualifications, personality fit, and motivation [31]. The primary goal remains unchanged: ensuring a mutually beneficial match between the individual and the organization. Achieving this requires a holistic evaluation of the applicant's data, ensuring they have the necessary prerequisite knowledge, can excel in the environment, and can fully utilize the resources offered by the organization or institution [31].

Additionally, It is worth mentioning that the application of the admission process extends beyond just education and professional fields, including areas such as healthcare, social services, and government programs. In these contexts, admission may involve evaluating medical eligibility, community support needs, or suitability for citizenship [18, 20]. Across all contexts, the objective is consistent—finding a fit that serves both parties well.

In light of the above, this thesis primarily investigates the core principles of the admission evaluation and decision-making process, focusing specifically on the context of higher education. This exploration is approached from the perspective of an **HCI (Human-Computer Interaction)** master's student, mainly focusing on the field of **Information Visualization (InfoViz)** and their role in improving user interaction with data and supporting informed decision-making.

In the following subsections, this introduction will provide a thorough examination of the higher education admissions process. Subsection 1.1 examines the influence of admissions on student life, discussing how the process shapes educational and personal outcomes. Subsection 1.2 focuses on the admissions process itself, highlighting the challenges and obstacles that institutions face in selecting the most suitable candidates. Subsection 1.3 discusses various strategies in aiding in the evaluation and decision-making processes, including Decision Support Systems (DSS) and Intelligent Decision Support Systems (IDSS), while also considering the human factors involved. Subsection 1.4 delves into the nature of human decision-making and evaluation models. Subsection 1.5 examines how admissions function as a Multi-Criteria Decision-Making (MCDM) process, highlighting the complexities of balancing various criteria. Subsection 1.6 shifts the focus to the field of data visualization, discussing its role in enhancing decision-making and multi-criteria evaluation processes, particularly in how visual tools can improve clarity and effectiveness. Subsection 1.7 reviews existing tools in higher education admissions, offering examples of data visualization techniques that have been integrated into the process while assessing their strengths, limitations, and areas for improvement. Finally, Subsection 1.8 presents the motivation behind this thesis, outlines its contributions to the field of Human-Computer Interaction (HCI), and introduces the research questions that guide the investigation.

1.1 The Influence of Higher Education Admissions on Student Life

The university admissions process is a crucial bridge that directly impacts students' academic and professional futures. The essays of child psychoanalyst *Erik Erikson*, which have significantly influenced the understanding of social development, argue that the period of young adulthood, which commonly intersects with the years spent in higher education, is a crucial time for forming one's personal and social identity [24]. Erikson suggests that young people's identity develops best when they are given a psychosocial moratorium: a diverse environment where young adults are granted the time and space to experiment with various roles, ideas, and career paths before committing to their long-term professional and personal goals [24].

The admissions process, therefore, is not merely a selection mechanism but a decisive factor in granting or denying young adults access to opportunities that influence their essential developmental stage [24]. It can either facilitate a student's journey of self-discovery and intellectual growth or, if insufficient, obstruct access to this critical exploration environment. Moreover, if the admissions process fails to adequately consider its environment's diversity and complexity, it could undermine its ability to function effectively as a psychosocial moratorium [24]. This underscores the substantial responsibility that higher education institutions bear to ensure that their admissions processes are fair and transparent.

1.2 Higher Education Admission: The Process, Challenges & Complexities

In terms of procedural steps, the higher education admission process typically begins with submitting applications, where students provide personal and academic records. Admissions committees then evaluate these applications, sometimes requiring additional input via interviews or entrance exams. Based on this evaluation, decisions are made to accept, waitlist, or reject applicants. Accepted students proceed to enrollment, which involves submitting additional documents and paying necessary fees. In cases where applicants believe their applications were not fairly assessed, objections may be raised, leading to a review or appeal process to ensure fairness and transparency. [59]

As globalization expands, the volume of submissions from diverse backgrounds has also risen, leading to an increase in the number of international students applying, a trend that continues to this day [15, 32]. [1]. [2]. [3]. [4].

[1]United States-based Common Application (CA) reported a 13% rise in international applicants for the United States college admissions for the past year, exceeding the 5% increase in domestic applicants [2]

[2]Nuffic, the Dutch equivalent of the CA, documented a 6.7% growth in new international students for the 2022-23 academic year [49, 48]

[3]In South Korea, the number of international students has steadily increased, reaching 181,842 in 2023, more than double the number 10 years ago (85,923), and ten times more than in 2004 (16,832) [73]

[4]In the United Kingdom, there were 679,970 overseas students studying at universities in 2021/22,

This surge in the number of applicants from diverse backgrounds significantly affects the entire admissions decision process and application evaluation, presenting several challenges. For example, a top student from the United States applying to a university in the Netherlands will encounter significant differences in grading systems and educational standards. The US grading system typically uses letter grades (A, B, C, etc.) and a GPA scale from 0 to 4.0, where an 'A' translates to a 4.0 GPA. In contrast, the Dutch system uses a numerical scale from 1 to 10, with 10 being the highest and grades above 8 considered excellent. This dissimilarity makes direct comparison challenging without understanding how different grades correspond to each other [75, 58].

Moreover, cultural differences in educational practices complicate the assessment of international applicants. Educational systems in different countries use varied approaches to assessing students, making comparing academic achievements from diverse backgrounds challenging [19, 32]. In addition to this, admissions officers must also consider the reputation, ranking, and academic standards of the institutions from which applicants come, as these factors help measure the quality of the applicant's previous education [32]. Hence, these differences in grading systems, assessment practices, and institutional credibility further complicate the comparison of academic achievements, thereby affecting the fairness and efficiency of admissions decisions [32].

Admissions committees assess a wide range of criteria, from objective data such as grades and test scores to subjective elements like personal statements and letters of recommendation [71]. Each component offers a unique perspective on the applicant, contributing to a holistic understanding of their potential [71]. These elements similarly stress the complexity of the situation, as universities worldwide—and even often within the same country—operate within different educational frameworks and standards, which further complicates the assessment, especially when the applicant's previous education differs significantly from the local institution's familiarities [15].

Addressing these dissimilarities in data interpretation across both international and, arguably, even domestic contexts, is crucial in the admissions process. While these differences can be cross-checked against international datasets and educational mapping tables to ensure alignment and accuracy [15], the complexity involved in evaluating diverse data points, especially when dealing with unfamiliar comparisons, cannot be overlooked. This process often requires a level of intuition and understanding that introduces a degree of uncertainty during the final decision-making stage, which can also lead to cognitive load and fatigue [32]. It is essential that this uncertainty is acknowledged and addressed to ensure that admissions decisions are efficient, fair, and reliable.

In further support of the above, examining recent acceptance rate data provides further insight into the selectivity, complexity, and critical importance of making well-informed admissions decisions in higher education, where the right call can offer life-changing opportunities [24]. For instance, Harvard's class of 2028—referring to students admitted in 2024 and expected to graduate in 2028—had an acceptance rate of just 3.6% [40]. Similarly, Stanford's class of 2026, admitted in 2022, reported an acceptance rate of 3.68% [7]. These low acceptance rates not only underscore the intense competition but also reflect the intricate nature of the admissions process at these institutions, where holistic, competency-based evaluations are employed [32, 59]. These evaluations consider both objective data, such as academic performance, and subjective assessments, like leadership potential and personal essays, making the decision-making process complex

including 120,140 from the European Union and 559,825 from other countries, marking a record total and a 37% increase over three years [55] and multifaceted [59]. Even at institutions with higher acceptance rates, such as Utrecht University [12], the challenge of balancing diversity and fairness with thoroughness in the admissions process is evident. It is crucial to recognize that the importance lies not just in the final decision but in understanding how each piece of data, from academic records to personal narratives, contributes to a comprehensive and transparent evaluation process and the outcome decision.

Beyond the aforementioned points, it is important to also reiterate that the diversity of data significantly amplifies the cognitive load on admissions officers [36]. Cognitive load refers to the mental effort required to process information and make decisions [30]. When admissions officers are already aligning numerous data points from applicants to form a comprehensive view, adding diverse and sometimes uncertain data that must be matched with institutional policies can become mentally overwhelming. The added strain can lead to fatigue and further complicate the decision-making process [34, 15].

Hence, to sum up, integrating all relevant factors into a cohesive final decision is essential. Whether decisions are made individually or through collaborative review, it is vital to recognize and address any inefficiencies. By fostering structured discussions where those with greater expertise or insight can effectively contribute, admissions committees can ensure a more thorough and balanced decision-making process. This collaborative approach not only enhances the accuracy of evaluations by incorporating diverse perspectives but also ensures that the decision-making process is transparent and well-documented, leading to more informed, equitable, and traceable outcomes [34].

1.3 Exploring Strategies to Aid Admission Evaluation & Decision-Making Processes

Building on the insights gained from the processes and challenges discussed in the previous subsection, the complex nature of university admissions decision-making has driven the exploration of various strategies and tools aimed at enhancing the fairness, efficiency, and transparency of this process. In particular, there has been a growing interest in the application of Decision Support Systems (DSS) and Intelligent Decision Support Systems (IDSS) [13, 57]. These systems, widely applied in fields like healthcare, finance, and logistics, offer structured, data-driven approaches to managing large volumes of complex data and suggest potential outcomes by automating portions of the decision-making process. Through algorithmic analysis, DSS and IDSS systems have shown their ability to handle multi-criteria decision-making, which aligns well with the needs of admissions processes, where both subjective and objective criteria must be evaluated [1, 13].

While these systems can automate many time-consuming tasks, such as data collection, administrative tasks, and preliminary scoring, they are not designed to completely replace human judgment. Human input is essential in interpreting DSS outputs and adding context-specific nuance to the evaluation [72]. For instance, an algorithm can assess grades and extracurricular activities, but it may struggle to fully capture qualitative attributes like leadership potential or personal fit for the program. Additionally, bias inherent in the training data of automated systems can propagate inaccuracies or unfair recommendations [21]. Therefore, while DSS tools provide significant efficiency gains, they have limitations when it comes to nuanced, holistic decision-making, which is critical in admissions.

At the intersection of DSS and visualization lies an important insight: these systems are most effective when their outputs are presented visually, aiding evaluators in interpreting complex datasets. Data visualization has become an indispensable tool in this context, acting as a bridge between raw data and human decision-making by reducing cognitive load and enabling more transparent, informed evaluations. In essence, visual representations such as charts and graphs help decision-makers synthesize vast amounts of information, facilitating clearer comparisons between applicants and enhancing the decision-making process [46, 27].

While DSS and IDSS offer useful frameworks and are worth mentioning for managing complexity, visualization tools provide a more transparent, intuitive means for admissions committees to assess and compare applicants holistically. By focusing on how data is presented visually—through charts, plots, and interactive dashboards—admissions officers can better grasp not only individual strengths and weaknesses but also the relationships between various data points. Therefore, the focus of this thesis shifts towards leveraging visualization as the primary mechanism to support holistic, transparent, and equitable decision-making in admissions, emphasizing the power of visual tools to enhance human judgment rather than replace it.

The next subsection delves deeper into how human decision-making works before exploring the role of visualization in supporting the admission evaluation process.

1.4 Human Decision-Making & Evaluation Models

Understanding the nature of human decision-making is crucial, particularly in complex contexts like university admissions. Before exploring how visualization can aid in these processes, it is essential to understand the cognitive mechanisms and psychological factors that drive how people think, evaluate things, and make decisions. Knowing these fundamentals helps in designing visualization tools that align with human thought processes, making them more effective in supporting evaluators during the decision-making process. By understanding how decisions are made, particularly in evaluating candidates, we can better tailor visual aids to enhance clarity and improve overall decision quality.

Herbert Simon's decision-making model, which divides the process into the phases of Intelligence, Design, and Choice, provides a structured framework that has been foundational in decision theory [63]. In the Intelligence Phase, decision-makers gather and analyze relevant information, identifying problems or opportunities. The Design Phase follows, where potential solutions are developed and evaluated against various criteria. Finally, the Choice Phase involves selecting the most appropriate solution from the alternatives. This methodical approach helps reduce errors and biases by ensuring that decisions are made with careful consideration of all relevant factors [50]. Importantly, evaluation is integral to the entire decision-making process, especially in the Design and Choice phases, as it underpins how alternatives are assessed and ultimately selected [50].

Recent studies suggest that the effectiveness of structured decision-making models is highly influenced by the ability of decision-makers to manage cognitive biases and psychological factors that shape judgment. For example, Oral et al. [50] highlights how cognitive biases can distort decision-making even when systematic approaches are used. They argue that biases, such as confirmation bias or anchoring, can interfere with an evaluator's ability to objectively assess information, particularly when dealing with complex datasets [50]. One way to mitigate these biases is through visualization tools, which present data in formats that are more intuitive and aligned with how humans naturally perceive and process information. Cibulski et al. [4] builds on this by demonstrating that effective visualizations not only simplify complex information but also clarify critical decision-making processes in domains where high cognitive load is common. Their work underscores the importance of visualization in facilitating more informed decisions by making the relationships between data points more transparent, tangible, and easier to interpret [4].

The complexity of decision-making in university admissions is further compounded by the need to balance multiple criteria, often with conflicting priorities. The European Union's Mastermind Europe Matrix [16] (see Figure 1) offers a methodological toolkit that harmonizes evaluation methods across international contexts, focusing on criteria such as subject-related knowledge, academic competencies, and language proficiency [16]. This toolkit provides a structured approach to evaluating applicants, emphasizing transparency and consistency in decision-making. It reflects Simon's model [63], where the Intelligence Phase involves gathering comprehensive data on applicants, the Design Phase involves evaluating this data against standardized criteria, and the Choice Phase involves selecting the most suitable candidates by accepting or rejecting the application.

How do you a) know (= assessment mechanisms) if b) students are good enough (= norms-levels) in c) the things they need to be good at.

Or, in logical order: 1) criteria, 2) norms/levels, 3) assessment mechanisms[6] with 4) testing scores

	1 Criteria	2 Norms/ levels	3 Assessment mechanisms	4 Assessment scores
	What you are looking for		What you are looking at	
Subject-Related Knowledge & Skills (Guiding tool 2)				
General Academic Competencies (Guiding tool 3)				
Personal Competencies & Traits (Guiding tool 4)				
Language competence (Guiding tool 5)				

Figure 1: The Coherent Admission Framework: the Matrix 1.0 [16]

Given these considerations, admissions inherently present a multi-criteria decision-making (MCDM) challenge. Admissions officers, as decision-makers, must go beyond simply analyzing data and evaluate how it aligns with various criteria while balancing conflicting institutional priorities [71]. Unlike analysts who focus on data quality and methodological rigor, decision-makers are responsible for converting this information into practical, strategic choices [50]. In this context, visualization tools play a pivotal role, helping decision-makers manage and balance these competing priorities by clearly presenting both qualitative and quantitative factors, allowing for more informed and comprehensive final decisions.

To better understand MCDM, the next subsection will delve deeper into the topic and explore how admissions function as an MCDM process.

1.5 Admissions as a Multi-Criteria Decision-Making Process

In university admissions, Multi-Criteria Decision-Making (MCDM) refers to evaluating applicants based on multiple criteria, each carrying varying levels of significance [59]. Admissions officers are responsible for assessing academic performance, personal qualities, extracurricular achievements, and the alignment of candidates with institutional goals [59, 71]. The complexity of balancing these often conflicting criteria is characteristic of MCDM's broader application in multi-faceted decision-making processes [71, 1].

MCDM serves as the overarching framework for making decisions involving multiple criteria. Within this framework, Multi-Criteria Assessment/Analysis (MCA) focuses on evaluating the relative importance of each criterion, often aligned with institutional goals and policy-driven priorities [9]. MCA allows decision-makers to compare alternatives across multiple dimensions without necessarily quantifying their significance [9]. Multi-Criteria Evaluation (MCE), on the other hand, is a more structured approach that incorporates numerical weights or scores to quantify the impact of each criterion [9]. By assigning weights, MCE provides a systematic and data-driven approach, allowing decision-makers to prioritize criteria based on their calculated significance, thereby supporting more objective and transparent evaluations [9] [3, 27].

In the broader context, Multi-Attribute Decision-Making (MADM) is closely related to MCDM, but it often focuses on selecting the best alternative from a set of options based on quantifiable attributes [9, 27]. This framework is particularly relevant in admissions, where alternatives (i.e., applicants) are evaluated against measurable criteria like grades, test scores, and leadership potential. Techniques such as the Analytical Hierarchy Process (AHP) and TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) are widely used in MADM for evaluating trade-offs between attributes to identify the most suitable alternative [80, 27]. These methods, which Dean [9] highlights as essential tools for formal decision-making processes, are used across various sectors, including education, to structure complex decision-making environments.

In multi-attribute choice tasks, decision-makers are tasked with selecting from a set of alternatives, with each alternative defined by multiple attributes [9]. In the admissions context, each applicant is considered an alternative, with attributes such as academic performance, leadership qualities, and extracurricular involvement being evaluated. Decision-makers must balance competing priorities and assess how well each alternative meets the defined criteria [71, 59].

Multidimensional visualizations play a significant role in supporting decision-makers during these tasks by presenting complex, multi-attribute data in an intuitive format [27]. Visualizations like parallel coordinates and radar charts enable decision-makers to explore multiple dimensions of the data simultaneously, making it easier to understand trade-offs and identify patterns [27]. Such tools are invaluable in reducing cognitive overload and fostering balanced, informed decisions. By presenting data visually, decision-makers can more effectively compare alternatives across multiple criteria, thus improving transparency and decision quality [53].

A major challenge in applying MCDM to admissions is the inherent subjectivity in evaluating certain criteria. Admissions officers may prioritize qualities like leadership or diversity differently, depending on personal biases or institutional goals. This subjectivity can result in inconsistent evaluations, especially when interpreting qualitative data such as personal statements or recommendation letters. Similar inconsistencies have been documented in other domains, like recruitment, where biases influence outcomes despite the presence of structured evaluation frameworks [27].

Uncertainty also plays a critical role in MCDM, as evaluators must assess not only an applicant's current qualifications but also their potential for future success. For example, the predictive value of standardized test scores or the long-term impact of extracurricular activities can vary significantly, making the decision-making process inherently uncertain. Addressing uncertainty is essential for ensuring fairness. Common approaches include assigning weights to overall scores based on the reliability or confidence levels of the data, a technique employed in fields such as risk management and investment analysis [52, 27].

To address these challenges, decision-support tools and visualization systems have proven increasingly valuable. These tools help transform complex, multidimensional data into more accessible formats, aiding admissions officers in understanding relationships between criteria and fostering more consistent, transparent decisions [27]. By visually representing trade-offs between criteria, these tools help reduce the risk of intuition-driven decisions influencing the final judgment, supporting more informed decision-making [27].

1.6 Field of Visualization & Its Role in Aiding Evaluation & Decision Processes

Visualization has become an indispensable tool for improving decision-making processes, particularly in environments that involve complex datasets or fragmented raw data that require synthesis to generate actionable insights [46]. The ability to translate abstract data into visual formats not only facilitates the identification of intricate patterns but also fosters deeper comprehension that might otherwise remain obscured in raw data [46].

A key advantage of visualization lies in its capacity to reduce the cognitive load associated with processing complex information [46]. By externalizing the cognitive processes, visualization allows decision-makers to focus on higher-order strategic thinking [44]. This capability is particularly valuable in high-stakes environments, where the swift and accurate processing of information is crucial for achieving optimal outcomes [14]. Visualization tools, acting as cognitive aids, simplify the complexities of data and facilitate the decision-makers ability to evaluate multiple criteria and consider trade-offs more effectively [46, 14].

Moreover, visualization plays a central role in mitigating biases and enhancing transparency in decision-making [46]. By structuring data into clear visual formats, these tools help anchor decisions in objective data, minimizing the influence of subjective interpretations that may arise from personal biases [8]. However, while visualizations reduce bias, they do not eliminate it. Human interaction with visual data can still be subject to cognitive biases such as confirmation bias and selective attention [11]. Poorly designed visualizations may even exacerbate these biases, drawing attention to irrelevant details or concealing critical information [11]. This challenge is especially prominent in complex decision-making environments, where both human cognitive limitations and visualization design choices can significantly impact outcomes. Thus, while well-crafted visualizations have the potential to simplify complex data, enhance transparency, and support cognitive processing, they must be carefully designed to avoid reinforcing biases or leading to misinterpretations, ensuring that they truly contribute to a more equitable and informed decision-making framework [46, 11].

Given these benefits, data visualization stands out as a promising approach for improving decision-making processes in contexts such as university admissions. The ability of visual tools to clarify data, reduce cognitive strain, and increase transparency aligns with the goals of promoting fairness and informed decision-making [14]. As the literature review will detail, exploring specific visualization tools and techniques reveals how these advantages can be effectively applied in practical scenarios, particularly within the admissions process in higher education. By focusing on the intersections of data complexity, decision support, and visual clarity, this research aims to identify best practices that can enhance both the operational and evaluative phases of the admissions process.

1.7 Evaluation of Tools in Higher Education Admissions: Examples of Data Visualization Integration

In higher education admissions, institutions often rely on a variety of in-house systems to manage application, evaluation, and admissions processes. The approach to these systems can differ notably, reflecting the diversity in educational frameworks and administrative practices worldwide [19, 42].

In the Netherlands, for example, the centralized system known as *Studielink* [67] streamlines the application process across multiple higher education institutions through a single, unified interface. This centralized approach simplifies the process for students while enabling institutions to manage applications more efficiently. However, within individual universities, the system operates in a decentralized manner. Each university employs its own internal system, often integrating with *Osiris* (example: `https://osiris.uu.nl`, `https://osiris.tue.nl`), a widely-used student information system in Dutch universities. Osiris handles various administrative tasks, including student records, course registration, grade reporting, and admissions.

Despite the utility of systems like Osiris in handling administrative tasks, their form-based nature and lack of advanced features for multi-dimensional data analysis or interactive visualizations highlight a gap in supporting more sophisticated aspects of the admissions process. In contrast, searching online platforms like *Capterra* `https://www.capterra.com` reveals a wide range of admission software tools designed to address these specific needs. A search for "admission software" on Capterra yields around 170 products offering features such as application management, activity dashboards, workflow management, and assessment scoring. For instance, tools like *Kira Talent* [69, 70] provide targeted features for the evaluation process, including AI detection scores and visualizations like parallel coordinate charts and simple bar charts to facilitate the evaluation of multiple advisors (see Figure 2).

Be that as it may, the majority of these tools focus on dashboard visualizations that provide an overview of institutional student data, such as demographics and trends, using basic visualizations like pie charts, bar charts, and time series graphs. As such, while these tools offer valuable insights into historical data and trends, they often fall short in supporting the intermediate stages of decision-making, where admissions committees must evaluate multiple criteria to reach a final decision. Critically, the uncertainty or metrics involved in final decision scores are often not visualized in a way that directly supports the decision-making process (see Figure 2).

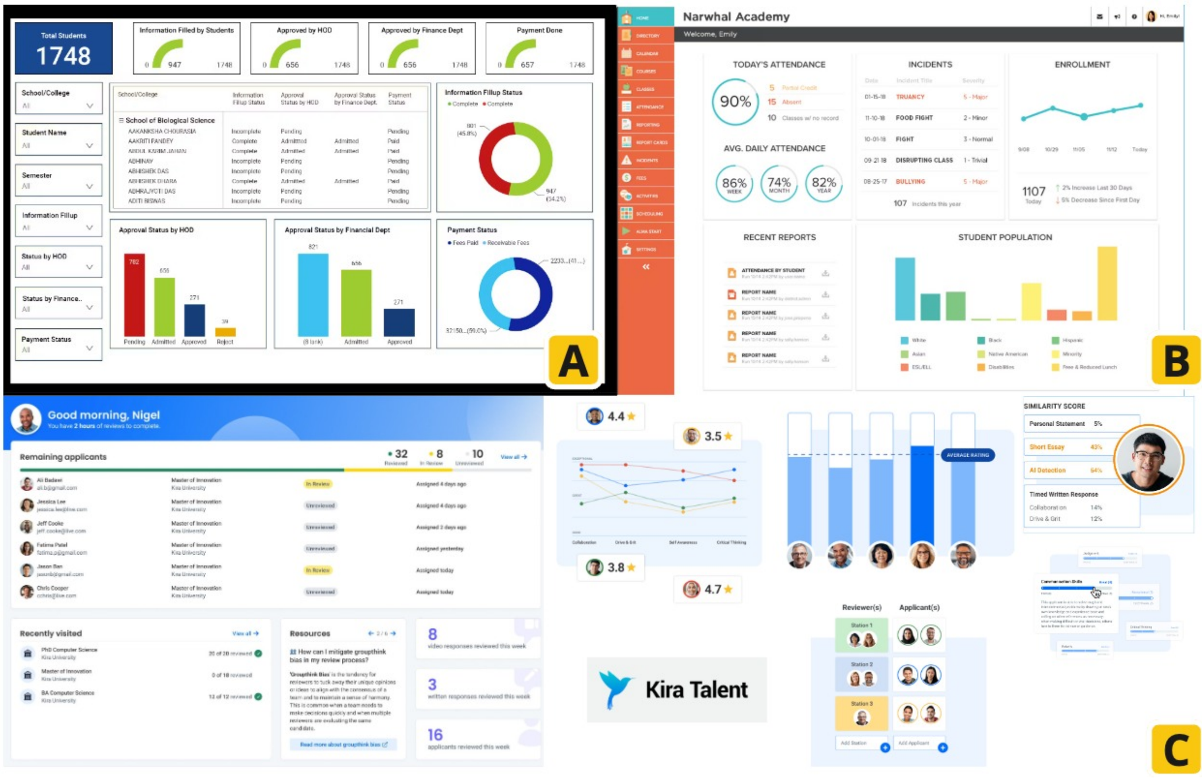


Figure 2: **Panel A** *MasterSoft* [65] features an administrative dashboard with pie charts and bar charts to track document approvals and financial statuses. **Panel B** *Alma SIS* [26] showcases student performance metrics such as attendance and demographic distributions, visualized through time series graphs and pie charts. **Panel C** *Kira Talent* [69] integrates evaluative tools, including radar charts and similarity scores for applicant comparison, though with a focus on AI-driven metrics. These tools commonly utilize educational data to inform strategic and operational decisions. However, they do not specifically address the nuanced evaluation process during admissions decisions—a gap this study aims to fill by introducing visualization tools designed to enhance multi-criteria evaluations and collaborative decision-making among multiple evaluators.

1.8 Research Questions & Thesis Contribution

Building on the insights from the previous subsections, there remains a gap in addressing the intermediate processes of university admissions, particularly in how individual data points are evaluated and influence the final decision. This gap is particularly noticeable in the use of visualization tools within admissions, where their potential to support collaborative decision-making has not been fully explored. The challenge lies in ensuring transparency and consistent evaluations while fostering collaboration among evaluators who bring diverse perspectives to the table.

To address the gap, this thesis investigates the role of data visualization in supporting the university admissions process, particularly in multi-evaluator environments, by posing the following research question:

Main Research Question:

How can data visualization be effectively employed to support the intermediate decision-making process in university admissions, particularly in evaluating individual aspects of an applicant's profile and facilitating collaborative assessments to form a holistic view for the final admission decision?

This main question is further broken down into four sub-questions.

Sub-questions:

RQ1: How do admissions committees make final decisions, and what factors are considered when evaluating applications?

RQ2: How can visualizations aid admissions committees in comprehensively assessing multiple criteria and making informed, transparent final decisions?

RQ3: How do visualizations help evaluators reflect on their own perspectives and those of their peers when assessing individual applicant aspects, and in what ways do these visualizations facilitate collaborative decision-making?

RQ4: How effective are visualizations in serving as conversation starters, reducing uncertainty, and reviewing the documented strengths and weaknesses of an application during the final decision-making process?

These research questions were formulated to address the challenges outlined in the problem statement. To develop a deeper understanding of how admissions committees evaluate applications in real-world scenarios and address RQ1, a pilot study was conducted through interviews with admissions experts at Utrecht University. This study focused on the first research question, examining the diverse data considered across various programs, ranging from interdisciplinary studies (which combine multiple academic disciplines) to more specialized fields. Building on these findings, the subsequent research questions were further explored by integrating insights from both the pilot study and the literature review. These insights informed the design and evaluation of the proposed visualization tool.

The evaluation process was conducted through a scenario-based approach and a simulated decision-making environment [35], where participants assumed evaluative roles within a controlled, admissions context. Custom datasets representing hypothetical applicants with varied application data were manually generated to simulate diverse applicant profiles and evaluation scenarios [35]. These controlled conditions allowed for an examination of how the proposed visualization tool could support evaluators in navigating multi-criteria decision-making tasks. By replicating real-world admissions activities within a structured environment, the study gathered insights into the tool's ability to facilitate more informed decision-making and collaboration among evaluators.

Through these simulations, the evaluation specifically addressed the research sub-questions while simultaneously contributing to the main research question, which seeks to explore how visualization tools can enhance the assessment of various aspects of an application and support arriving at a final, well-informed decision.

1.9 Structure of the Thesis

This introduction has established the foundation for the thesis by outlining the key issues and motivations behind the research. It has provided an overview of the domain description, introduced the research questions, and presented the planned approach for addressing these questions.

The following chapters now build upon and explore the theoretical basis of this thesis. Section 2 presents a review of the relevant literature on information visualization in decision-making and collaborative decision-making. Section 3 describes the pilot study conducted to gain insights into the needs and challenges faced by admissions officers in real-world scenarios. It details the methodology, data gathered, and key findings that informed the development of the proposed visualization tool. Section 4 introduces the proposed tool, ***EvaluationViz***, designed to address the challenges in university admissions during the evaluation phase. The section will outline the tool's design, features, and the rationale behind its development, connecting the study to the field of information visualization. Section 5 covers the evaluation methodology and presents findings from the assessment of EvaluationViz. Section 6 is dedicated to discussing the implications of the findings, addressing study limitations, and proposing areas for future research. Finally, Section 7 concludes the thesis by summarizing the research contributions, reflecting on the findings, and offering final thoughts.

2 Literature Review

This literature review synthesizes insights from several key studies within the context of information visualization and the application of visualization frameworks in multi-criteria decision-making (MCDM) and group decision-making processes. Additionally, related studies that evaluated visualization approaches in similar admission or evaluation decision-making contexts are considered. This section provides a foundation for understanding how these principles can serve as a knowledge base and be further extended to develop a creative solution tailored to address the research question outlined in this thesis. [5]

2.1 Visualization Frameworks for Data-Driven Decision-Making

In complex decision-making environments like university admissions, where multiple criteria are considered, visualization has the potential to play a significant role in organizing and synthesizing data for evaluators. To achieve this, frameworks such as Tamara Munzner's Nested Model for Visualization Design [45] provide valuable guidance. This model (see Figure 3), structured around four key stages—task abstraction, data abstraction, visual encoding, and iterative refinement [45]—offers a systematic approach to the design process. By focusing on how to represent data in ways that facilitate accurate and effective interpretation, Munzner's framework can be valuable guidance in the journey of making solutions for aiding admission evaluation using visualization [46].

[5] Several primary search terms used to gather relevant literature included "Information Visualization," "University Admission and Human-Computer Interaction," "Visualization in Admission or Evaluation Processes," "Multi-Criteria Decision Making and Visualization," and "Collaborative Decision Making and Visualization."

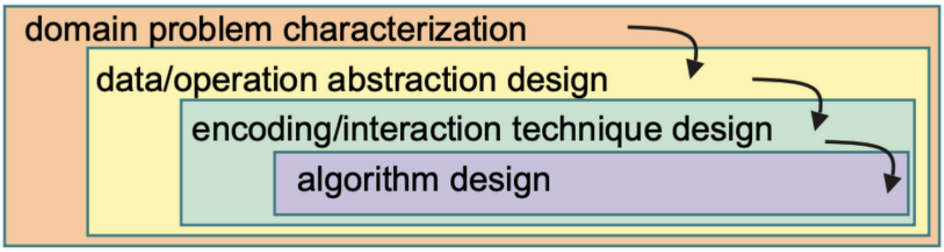


Figure 3: Tamara Munznerś Nested Model [45]

In the context of university admissions, task abstraction involves breaking down the complex evaluation process into specific, manageable tasks. As already addressed, these tasks typically include assessing academic performance, reviewing personal statements, and evaluating other relevant attributes of an applicant. By focusing on these discrete components, visualizations can potentially guide admissions committees toward a more holistic assessment of each applicant's profile, encompassing all dimensions [71, 45]. This exploratory approach suggests that visualizations could support both detailed and high-level evaluations, enabling more comprehensive decision-making [46].

Once task abstraction is complete, the process moves towards data abstraction, which refers to organizing raw applicant data into structured formats suitable for visual representation [45]. In university admissions, this can be interpreted as involving organizing quantitative data—such as grades and test scores—while also providing a framework for qualitative judgments, like how well an applicant aligns with program goals. Structured data abstraction offers the possibility of applying consistent metrics to each data point, which may assist evaluators in more clearly assessing multiple dimensions of an applicant's profile. This approach can present the various aspects of each applicant cohesively, potentially improving clarity and reducing complexity for decision-makers.

Following data abstraction, the next critical step is visual encoding, which involves mapping structured data onto visual elements such as position, size, and color [45]. This process is crucial in determining how effectively evaluators can interpret the data [46]. Various visualization techniques can be employed depending on the nature of the data. For example, bar charts, which use length-based encoding, have been shown to be more perceptually accurate than area-based representations like pie charts or bubble plots [60, 68]. This makes bar charts particularly effective for representing quantitative data, where precise value comparisons are necessary [68]. When comparing multiple criteria, parallel coordinates can be useful, as they can reveal relationships between different evaluation metrics or highlight variations between evaluators' judgments [29]. On the other hand, stacked bar charts allow for the combination of various aspects of an applicant's profile into a single visualization. However, caution is advised with stacked bar charts due to potential perceptual challenges, particularly when comparing areas that lack a common baseline [68]. Another option, scatter plots, can be particularly effective for identifying patterns or outliers within an applicant's data, offering insights across multiple dimensions. Still, scatter plots require careful design to maintain readability. Ensuring clarity involves avoiding both visual clutter and excessive white space, which can otherwise result in insufficient data being displayed [61]. These visualization techniques offer various strengths and challenges, all of which should be carefully considered when designing a solution that effectively addresses the research question.

When exploring possible algorithms to analyze applicant data, Ultsch and Lötsch [74] has shown techniques such as Euclidean distance can be used to measure the similarity or difference between applicants based on multiple criteria. This method calculates the straight-line distance between data points in a multi-dimensional space [74], making it useful for being utilized as comparing applicant profiles or evaluating how closely aligned different evaluators' judgments are. It can also be applied to compare applicants directly to one another. Once these distances are calculated, they can be visualized using methods such as scatter plots or line charts, which help reveal patterns, clusters, or outliers in the data [78]. By incorporating Euclidean distance into visualization algorithms, it becomes possible to map the degree of similarity or divergence across multiple dimensions, which offers insights into how applicants compare across various criteria or how evaluators rank them [74]. While typically used in clustering and classification tasks, Euclidean distance can be considered as a potential algorithm to provide a quantifiable measure that can enhance visual representations and support more informed decision-making.

Building on this, it is essential to balance presenting these insights with minimizing cognitive load for the user. Well-designed visualizations must avoid clutter and prioritize readability, ensuring that admissions committees can efficiently interpret the information at a glance [11]. Striking this balance is particularly important in high-stakes decision-making environments, where accurate and swift evaluations are required.

In line with this, Cleveland [5] research on the perceptual accuracy of different display methods is particularly relevant when designing visualizations for admissions. His work demonstrates that design choices, such as the use of grid lines, aspect ratios, or reference frames, can significantly influence how data is interpreted [5]. These design considerations are critical in preventing misinterpretations that could lead to inaccurate assessments of an applicant's qualifications.

In designing visualizations for admissions committees, it is also important to account for the cognitive characteristics and expertise of the users. As highlighted by Conati et al. [6], decision-makers cognitive styles and familiarity with visualization tools affect how effectively they engage with the data [6]. Given that admissions committees often consist of individuals with varying levels of experience and technical skill, visualizations must be designed to be accessible and interpretable for all users, ensuring that each member can engage with the data and make informed decisions [34].

Beyond static visualizations, interactive visualizations, as explored by Perdana et al. [56], add a valuable layer of flexibility by allowing users to explore data dynamically [56]. In collaborative decision-making settings like admissions, interactive tools enable committee members to adjust parameters, explore different data dimensions, and view applicant profiles from multiple perspectives. This interactivity has the potential to foster a more comprehensive understanding of each applicant and support more informed, collective decision-making.

These theoretical insights have directly informed the design of EvaluationViz, the visualization tool developed in this thesis. The principles of task and data abstraction, visual encoding, and user-centered design have guided the creation of a tool that meets the specific decision-making challenges of university admissions, ensuring that it aligns with the research question and supports both individual and collaborative evaluations.

2.2 State of the art in Multi-Criteria Assessment, Weight-Based Decision-Making, and Group Collaborative Decision-Making visualization

In the previous subsection, basic visualizations were covered. This subsection shifts the focus to more composite visualizations—those that build on fundamental visualization techniques by integrating more advanced features, such as group collaboration and multi-criteria decision-making support. These tools often combine various visual encodings, allowing users to interact dynamically with data. By doing so, they enhance decision-making in environments that require the balancing of multiple factors. The visualizations reviewed in this section include multi-criteria and weight-based adjustments and group preference mechanisms, and they aim to facilitate more informed, transparent decisions.

One such visualization is **ValueCharts** [3] (see Figure 4), a composite visualization framework designed to facilitate group decision-making. It enables users to visually inspect and adjust preferences based on conflicting objectives [3]. ValueCharts simplifies the analysis of linear models of preferences, where decision-makers must often navigate trade-offs between multiple criteria [3]. The tool uses stacked bar charts as a primary encoding method to display the cumulative impact of each decision criterion [3, 27]. By visualizing how criteria like cost, performance, or other attributes contribute to the total value of an alternative, decision-makers can easily grasp the relative importance of each factor [3]. Beyond stacked bar charts, ValueCharts allows for weight-based adjustments, where users can modify the weight of each criterion in real-time, instantly visualizing the impact on overall rankings [3]. Although ValueCharts offers flexibility and interactive features that are beneficial in group settings, its primary focus is on providing an aggregate view across multiple options, such as comparing candidates in a ranking. However, in admissions, decision-makers often need to evaluate individual candidates across multiple dimensions in a straightforward and intuitive way. ValueCharts is designed to visualize trade-offs between multiple alternatives, but it may not fully address the need for a more focused, simplified, and detailed analysis of individual candidates through visual encoding.

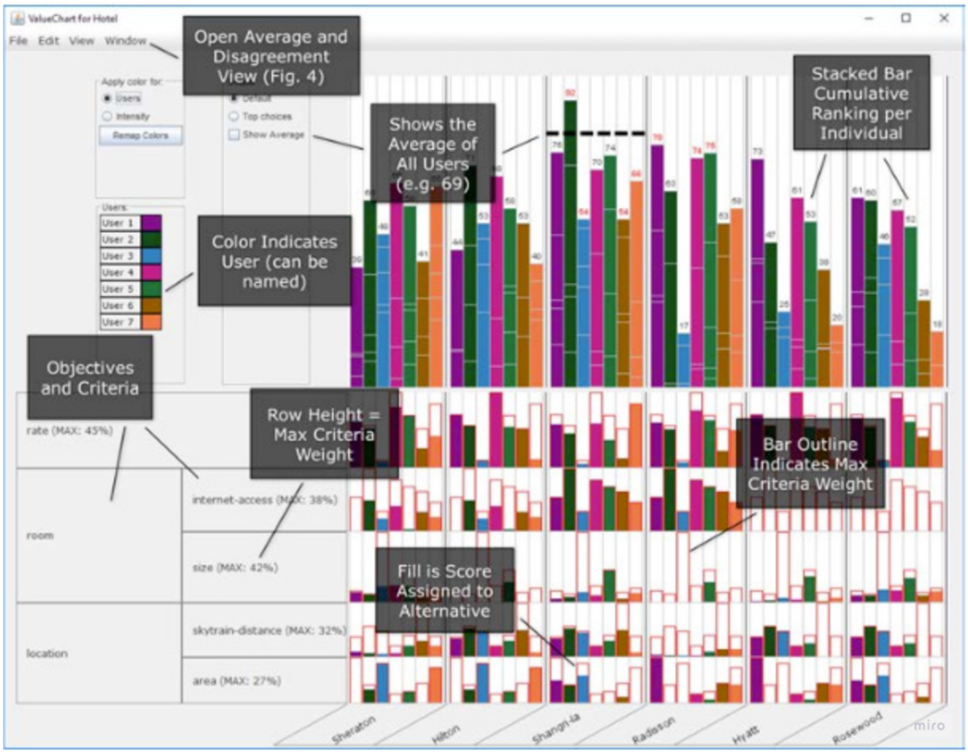


Figure 4: An example of ValueChart displaying the aggregated preferences of multiple users. The stacked bar chart represents different evaluators' rankings of criteria, showing both individual and group preferences. The tool allows users to adjust the weight of each criterion dynamically, providing real-time feedback on how these adjustments impact overall rankings. Color indicates user identity, and the bar height corresponds to criteria weight, while row height represents maximum criteria weight across users. This visual framework supports interactive, multi-criteria decision-making and enables users to explore trade-offs between criteria (extracted from [3]).

Similarly, **WeightLifter** [53] (see Figure 5) offers a focused approach to weight-based decision-making, particularly in scenarios where dynamic adjustments to criteria are necessary [53]. While it is not specifically designed for admissions, WeightLifter emphasizes real-time interaction with data, allowing users to modify the importance of specific criteria and immediately see how these changes impact rankings [53]. The tool uses visual encodings like bar charts and radar charts to display how changes in weighting various factors affect the overall evaluation [53]. This makes WeightLifter particularly well-suited for contexts where decision-makers need to explore multiple weighting scenarios to understand how shifting priorities can influence outcomes. WeightLifter allows users to interact with up to 10 criteria, adjusting weights and visualizing decision sensitivities [53]. However, admissions processes may involve more than 10 criteria, requiring flexibility. Additionally, the complexity of navigating weight spaces in WeightLifter might not suit the need for simple, intuitive visualizations that allow admissions officers to quickly grasp key insights without extensive interaction or adjustments. Therefore, while WeightLifter seems well-suited for contexts such as policy evaluation, particularly for analyzing the outcomes and fallout of students using educational data, or evaluating the policy itself, it does not fully support the need for case-by-case candidate evaluation.

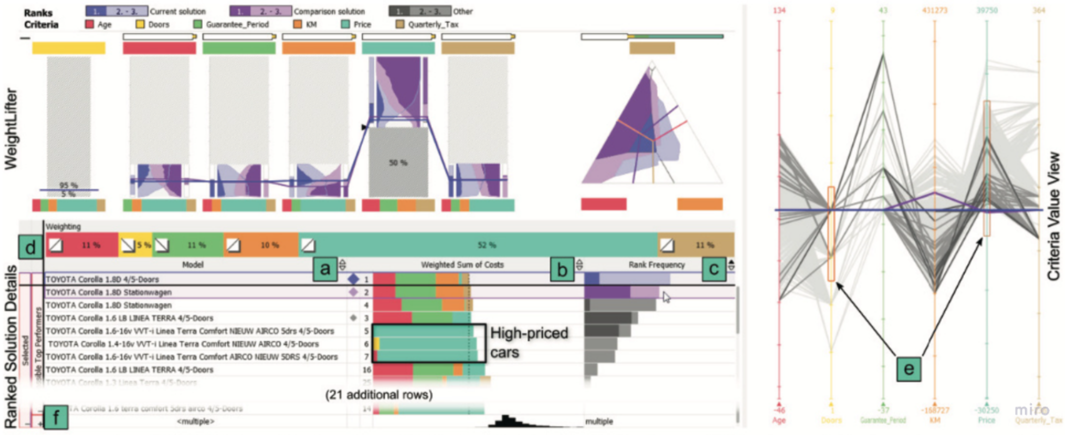


Figure 5: WeightLifter's interface allows users to dynamically adjust the weights of different criteria and visualize the impact on decision rankings in real-time. It employs a combination of bar charts and radar charts to display how changes in weighting affect the overall rankings of alternatives. This tool is designed to support weight-based decision-making processes, providing a flexible and interactive method for evaluating how priority shifts alter decision outcomes (extracted from [53]).

Building on this, Hindalong et al. [28] explores the challenge of designing preference visualization tools that balance simplicity with depth. Hindalong et al. [28] emphasizes that group decision-making visualizations should not only allow for composite visual encodings like parallel coordinates and bar charts to represent individual preferences but also facilitate the discovery of consensus while exploring individual differences within the group [28]. This capability is particularly important when decision-makers must reconcile differing priorities in a way that supports transparent, informed discussions without overwhelming users with excessive complexity [28]. Such tools are crucial when seeking to bridge individual preferences with collective decisions in diverse, collaborative environments. Although what Hindalong et al. [28] emphasizes is important to acknowledge, the education evaluation process is not straightforward, as revealed in the pilot study. While identifying differences in data inputs from advisors is valuable, it is crucial to strike a balance so that the decision is not unduly influenced by visualization but rather by an open conversation among evaluators to understand the variations in their inputs. Individuals come from diverse backgrounds and possess varying levels of familiarity, especially in multidisciplinary programs where the admissions committee oversees the overall evaluation. They also seek input from experts familiar with specific candidate backgrounds, as interdisciplinary courses often encompass a wide variety of expertise. However, this input should not disproportionately influence the overall decision of the admissions committee, which should focus only on relevant aspects. The goal remains to gather and combine inputs without over-manipulating or biasing the decision through visualizations that might skew outcomes.

Building on the advantages of flexible visualizations, **LineUp** [23] (see Figure 6) offers a powerful framework for creating, exploring, and refining multi-attribute rankings by combining attributes from heterogeneous datasets and visualizing how changes in these attributes impact overall rankings [23]. One of its key features is the ability to track ranking changes over time, which makes it especially useful for evaluating candidates across different years, where policy updates might occur. However, during the evaluation of candidates in a specific year, even if the process spans multiple days, the policy must remain consistent to ensure fairness across the group. In these cases, LineUp's functionality to experiment with weights and criteria over time is more relevant when looking at trends in historical data or understanding the impact of policy shifts, but less suited for situations where real-time adjustments would compromise fairness within a single admissions cycle [23].

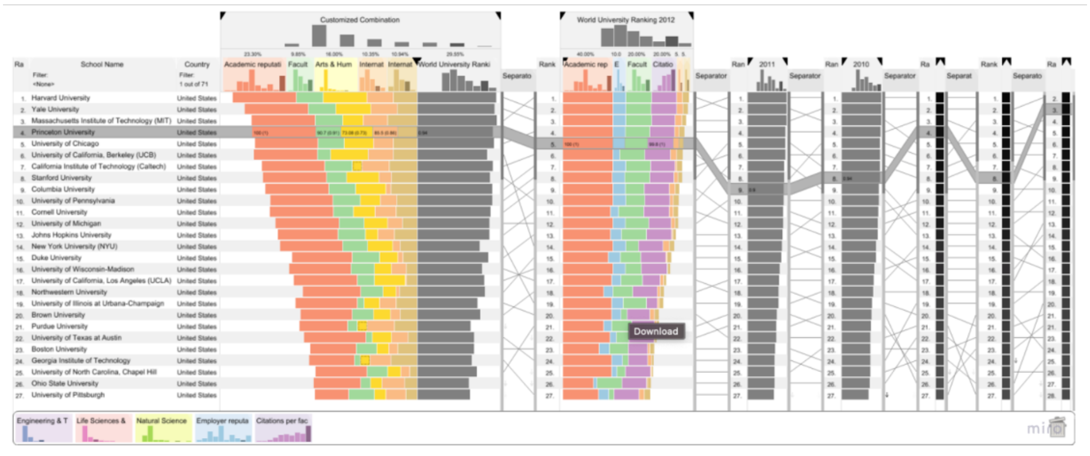


Figure 6: A visualization from LineUp, demonstrating how multiple attributes are combined to create rankings. The stacked bar chart shows the contribution of each attribute to the final ranking, while the slope graph tracks how rankings change over time (extracted from Gratzl et al. [23]).

Building on this, **ConsensUs** [41] (see Figure 7) is another visualization tool that focuses on facilitating group decision-making by highlighting areas of agreement and disagreement among evaluators [41]. Unlike the quantitative focus of tools like ValueCharts and WeightLifter, ConsensUs promotes collaboration by using dot plots and small multiples to compare individual preferences against group averages. This approach is particularly useful for fostering transparency in scenarios such as admissions, where multiple evaluators need to reconcile differing opinions [41, 27]. However, further research is needed to explore how ConsensUs can be adapted to ensure that the collaborative nature of its visualizations does not inadvertently shift focus. Although it can be incorporated as one visualization view, it is essential to enable evaluators to examine all aspects and dimensions, including visualizing the metrics involved in each decision, such as the uncertainty contributing to evaluating each application data point.

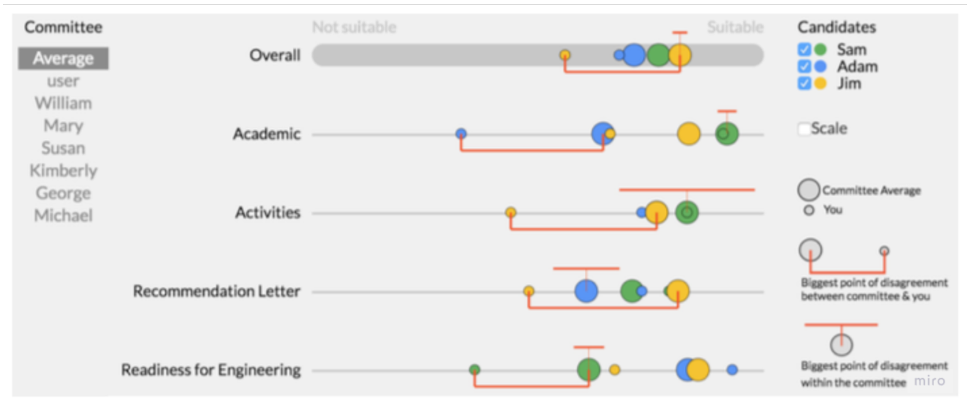


Figure 7: ConsensUs provides a visualization platform designed for group decision-making by allowing users to compare their individual rankings with group consensus. Through the use of dot plots and small multiples, it reveals areas of agreement and divergence among evaluators. This tool emphasizes transparency in collaborative decision-making by making individual rankings visible alongside the group's collective decision, encouraging discussion and alignment (extracted from [41]).

Research by Dimara et al. [10] provides valuable insights into the methodological challenges of evaluating visualizations in decision-support systems. The study emphasizes the importance of rigorously assessing how visualizations influence decision-making processes. It highlights the risks associated with perceptual biases in visualization design and underscores the need for simplicity and clarity in visual representations [10]. For example, tabular formats, which are often viewed as less sophisticated, can outperform more complex visualizations in supporting fast and accurate decision-making in scenarios that require evaluators to process multiple dimensions of data quickly [10].

In addition to ranking tools, another visualization technique, which is also relevant in the context of students and the university domain, is found in Heileman et al. [25]. Heileman et al. [25] introduces a Sankey diagram-based approach to understanding how students progress through educational systems [25]. These visualizations help decision-makers track student progress over time, uncovering patterns in enrollment, dropouts, and graduation rates. By displaying the flow of student populations through various academic stages, Sankey diagrams offer a high-level overview that is invaluable for identifying trends and making data-driven decisions about student success [25]. While this technique is primarily used in educational administration, its approach to visualizing transitions and progress could inspire future adaptations for admissions decisions. Visualizing the "flow" of applicants through stages of evaluation or how their profiles evolve with new information could provide a similarly intuitive way to track decision-making progress.

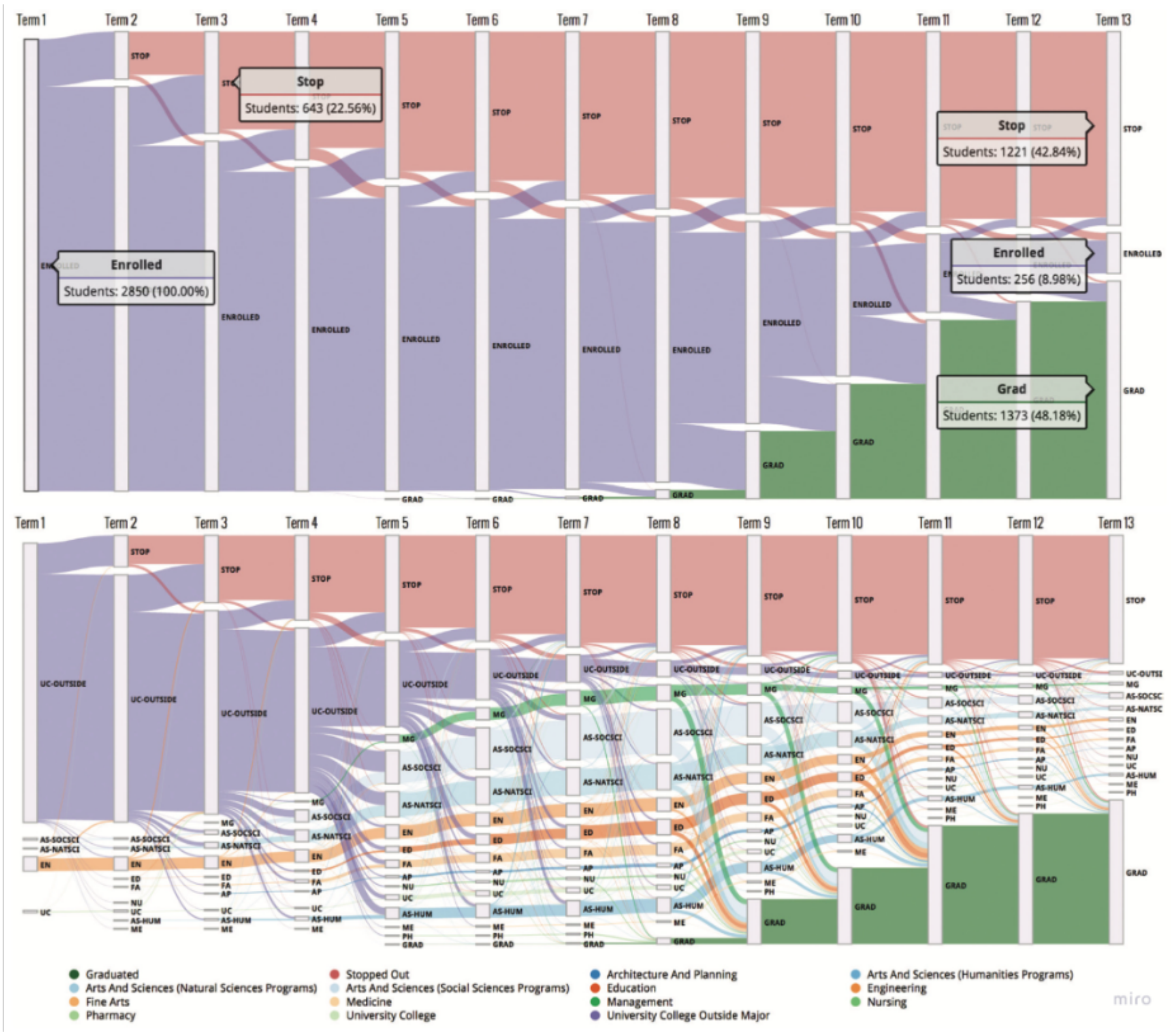


Figure 8: A Sankey diagram illustrating the flow of students through different academic stages, helping visualize enrollment trends, dropouts, and graduation outcomes over time (extracted from [25]).

In addition, it is important to consider visualizations that have explored qualitative data and can be mapped in the context of assisting in the evaluation of subjective data, such as motivation letters, in university admissions. Admissions evaluators often interpret qualitative inputs differently, depending on their personal priorities and interpretations. Here, visualizations can play a vital role by synthesizing key insights highlighted by each evaluator. For instance, word clouds can visually represent the most important themes across evaluators, making it easier to identify recurring strengths or concerns in an applicant's letter. However, word clouds also present limitations, such as randomization and lack of context, which can distort the true importance of terms if not used carefully. Addressing these issues through structured layouts or weighted frequency distributions, such as those proposed in Lohmann et al. [39], could improve accuracy and mitigate the effects of randomization typically associated with word clouds [33, 39]. Additionally, sentiment analysis visualizations could reflect the tone of each evaluator's assessment of qualitative content, using color-coded sentiment bars or heat maps to show how positive, neutral, or negative the letter was perceived [33].

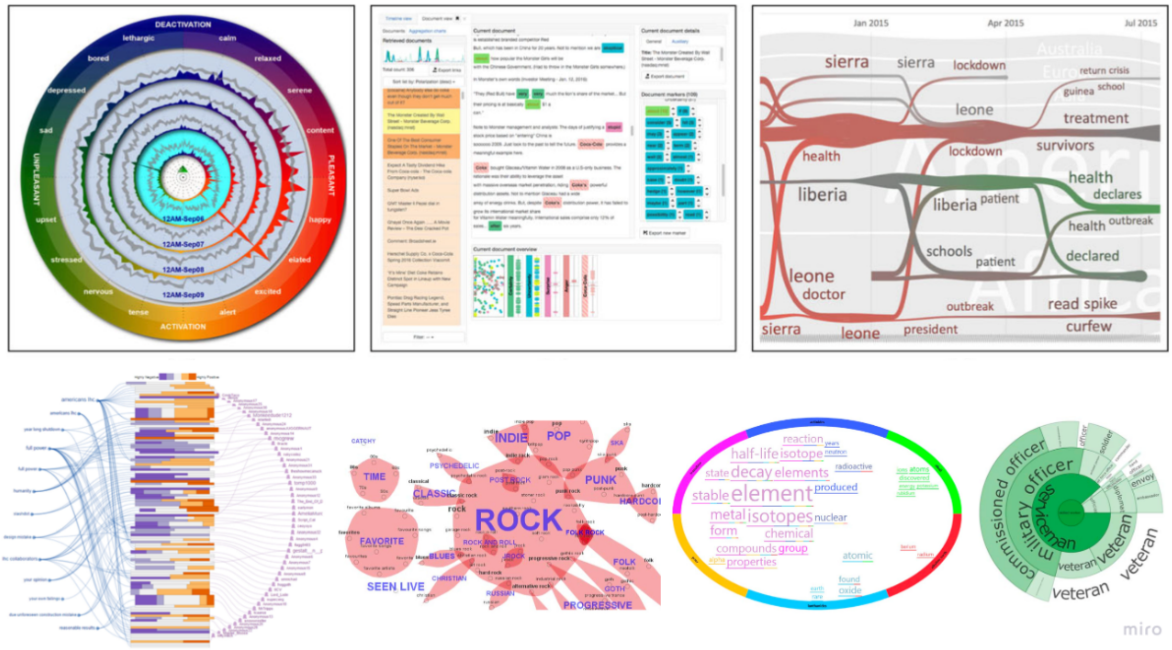


Figure 9: This combined figure showcases both word cloud and sentiment analysis visualizations. The word cloud illustrates the frequency of key themes across evaluators' feedback, with larger words representing higher occurrence. The sentiment analysis heatmap displays the emotional tone of qualitative evaluations, using color-coded bars to indicate positive, neutral, or negative sentiment. These visualizations inspire insight into subjective data interpretation in admissions processes. More inspiration examples can be found at `https://textvis.lnu.se/`. (extracted from [33, 39]).

The listed visualization papers have been evaluated to perform well in their respective contexts and are also potential tools for exploring both objective and subjective assessments at an individual level and in collaborative settings. These visualizations offer potential in the admissions context, as they not only assist in understanding subjective data at an individual level but also provide a collaborative view of how different evaluators' qualitative inputs align or differ. By integrating evaluators' comments and highlighting relevant aspects in a shared visual space, decision-makers can easily compare different perspectives before reaching a final judgment. However, these visualizations must preserve the depth of subjective insights without oversimplifying the narrative, thus maintaining their richness and meaning. While the mentioned tools support criteria adjustment, ranking visualization, and collaboration, further research is necessary to optimize their use in admissions decision-making. The focus must remain on simplicity and clarity to ensure that evaluators can quickly interpret key insights without introducing bias or confusion. For example, one could argue that using multiple bar charts side by side may not effectively highlight the suitability of an applicant for the program and could lead to information overload, making it difficult to compare all the data at once. Additionally, allowing evaluators to view metrics like weight alongside uncertainty would enhance the understanding of the subjective data. A gap for further exploration is how these tools can support case-by-case evaluations, emphasizing individual fit rather than merely ranking candidates and selecting the best fit, while considering all decision-making dimensions such as uncertainty and the relative weight of different factors.

2.3 Metrics and Dimensions Influencing Decision-Making

In decision-making processes, especially in subjective and multi-criteria evaluations such as university admissions, it is crucial to consider the factors contributing to each decision point. These factors impact the transparency and fairness of the process and influence the confidence and accuracy of the final decisions. Uncertainty plays a key role in situations where clear numerical evaluation is impossible, such as when assessing personal statements or motivation letters. According to Padilla et al. [52], there are multiple types of uncertainty that need to be visualized to support better decision-making. These include aleatoric uncertainty, which refers to inherent randomness, and epistemic uncertainty, arising from a lack of knowledge [52]. Visualizing uncertainty through techniques like error bars, violin plots, and ensemble displays helps decision-makers understand the variability in subjective judgments, reducing the risk of cognitive biases such as deterministic construal errors [52].

Dimara et al. [11] further emphasizes that cognitive biases often arise from misjudging data certainty, particularly when decision-makers misinterpret visualizations. They propose a task-based taxonomy that highlights the importance of using uncertainty visualizations to mitigate misinterpretation and bias. For example, interval plots and probability density functions can reduce bias by representing the range of possible interpretations, making decision-makers more aware of the underlying uncertainty in the data [11].

Another essential dimension is the weight assigned to various criteria in admissions decisions. Studies such as Pajer et al. [53], addressed in detail in the previous subsection, emphasize the role of interactive visualizations in dynamically adjusting these weights and observing their impact on overall rankings. However, such tools must avoid overwhelming users with complexity, as decision fatigue can negatively affect outcomes.

The visual encoding used to represent weights and uncertainty is critical for enhancing decision-making efficiency, allowing users to differentiate between various elements without cognitive overload.

In conclusion, visualizing all dimensions—such as uncertainty and weight—that influence decision-making is essential to ensuring fairness, transparency, and confidence in outcomes. Tools that effectively balance these dimensions with clear visual distinctions offer the potential for improving admissions processes by reducing bias and increasing decision accuracy, which is a point to be considered during the development of the visualization solution for this thesis.

2.4 Bias-Aware and Interaction-Analytical Behavior-Aware Visualization Systems for Decision-Making

In the realm of decision-making, particularly in domains such as candidate selection or university admissions, bias mitigation and interaction awareness have gained significant attention. Recent advancements in visualization tools like **Lumos** [47] and **BiasEye** [38] present novel approaches to tackling unconscious biases and promoting transparent decision-making processes.

Lumos [47] (see Figure 10) is a visualization tool aimed at increasing users' awareness of their analytic behaviors during data exploration. By capturing interaction histories both in-situ (directly within the user's data analysis) and ex-situ (through external comparison views), the system provides real-time feedback to users, making them more conscious of how their focus might be shaping their decisions [47]. Lumos is particularly valuable in revealing biases in decision-making by tracking which data points or attributes are emphasized or overlooked [47].

To achieve this, Lumos uses a range of visual encodings, such as bar charts and scatter plots, to display interaction traces overlaid on the data distributions [47]. By visually representing user interactions with color gradients, the tool highlights which data points have been interacted with frequently and which have been neglected, offering a clear view of how focus may skew decision-making. For example, a dark blue color in a scatter plot indicates a higher level of interaction with certain data points, while lighter colors signal areas that have received less attention [47]. These visual encodings help users identify patterns of overemphasis or neglect, fostering more balanced and informed decision-making. The system's ability to visualize interaction traces in real-time enables users to adjust their focus and reflect on potential biases during the data exploration process [47].

Figure 10: The Lumos interface consists of seven key sections, including traditional data analysis panels and interactive visualizations. It displays analytic behavior through both in-situ and ex-situ interaction traces, helping users compare their focus with underlying data distributions across various visual elements (extracted from [47]).

Similarly, **BiasEye** [38] (see Figure 11) focuses on mitigating cognitive biases in material screening processes, such as recruitment and admissions. BiasEye integrates bias-aware design into its visualization system by using machine learning to model individual screening preferences and biases [38]. The system uses a variety of visual encodings, including box plots, scatter plots, and glyphs, to represent statistical distributions and human scores alongside model predictions [38]. A mixed-design user study demonstrated that BiasEye increased participants' awareness of potential biases, such as anchoring and screening order effects, which can significantly influence decision outcomes [38]. By modeling individual screening preferences, BiasEye can flag instances where certain attributes, like the order of candidate review, disproportionately influence the final decision [38]. This feedback loop allows evaluators to adjust their screening practices, ensuring a more impartial assessment process. BiasEye's success in improving decision-makers' confidence and consistency highlights the potential of bias-aware systems in critical decision-making environments [38].

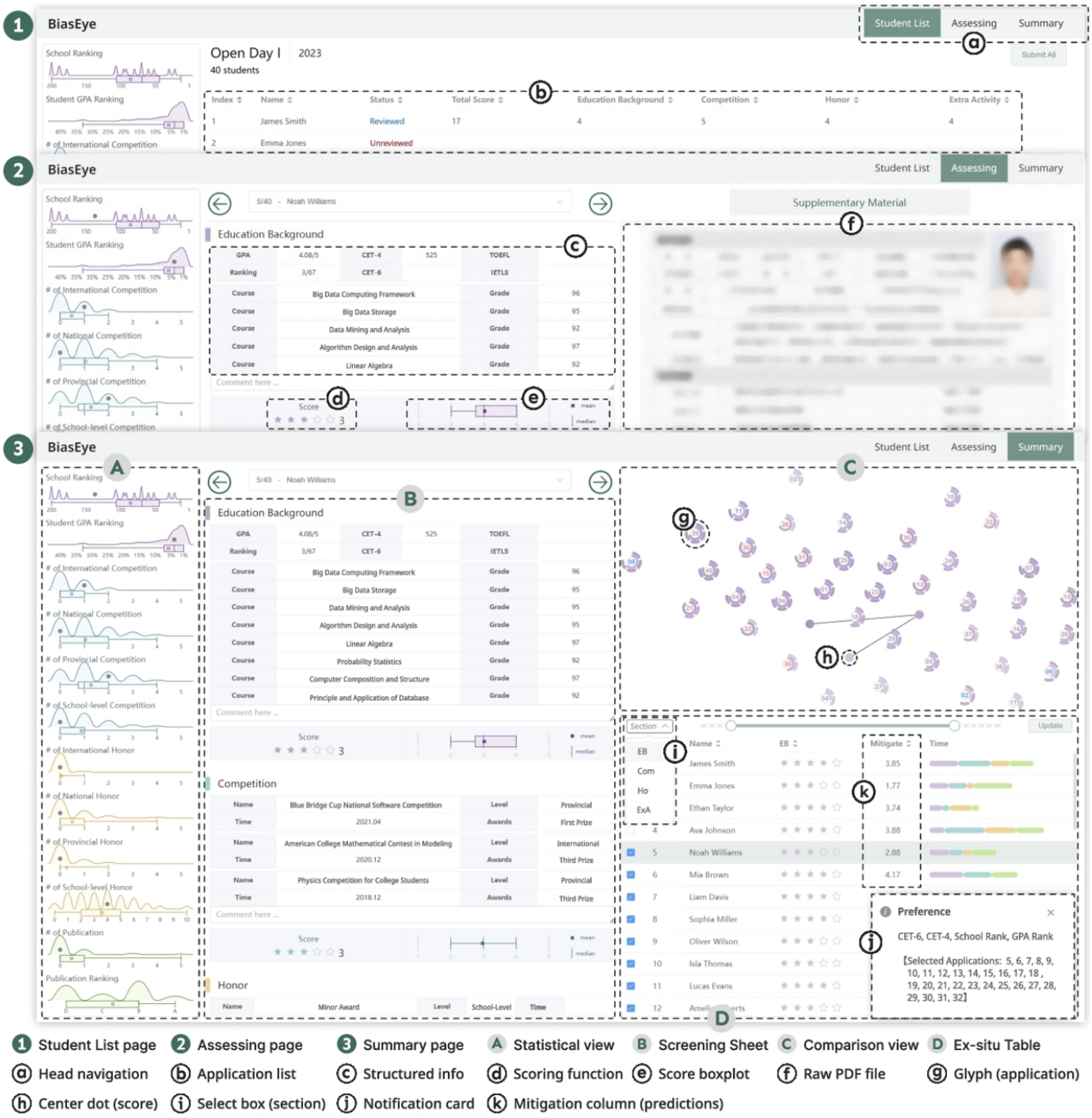


Figure 11: The frontend design of BiasEye, showcasing its bias-aware visualization system with various visual encodings, including box plots, scatter plots, and glyphs, used to model and mitigate cognitive biases during the screening process (extracted from [38]).

Both Lumos and BiasEye emphasize the importance of transparency and accountability in visual data analysis and decision-making, particularly in settings like admissions and recruitment. Their approaches highlight the value of real-time interaction feedback and bias detection, which are critical for ensuring fair and impartial evaluations [47, 38]. By integrating bias awareness and analytic traceability, these tools contribute to more balanced, data-driven outcomes [47, 38]. However, while Lumos focuses on tracking user interactions to raise awareness of analytic biases [47], and BiasEye targets the identification and correction of cognitive biases in screening processes [38], this thesis takes a broader, more holistic approach. The goal here is to provide a comprehensive view for evaluating decisions across multiple dimensions, rather than concentrating solely on bias reduction.

This expanded focus allows for a more complete understanding of the factors influencing decision-making, ensuring that all relevant aspects are considered in a transparent and balanced manner.

2.5 Conclusion of Literature Review

The reviewed literature underscores the pivotal role of visualization and interactive tools in improving multi-criteria decision-making (MCDM) for university admissions. Munzner's Nested Model of Visualization [45, 46] offers a strong theoretical foundation for designing visualizations that enhance clarity and user interpretability. Building on this, tools such as ValueCharts [3] and WeightLifter [53] demonstrate the value of dynamic weight adjustment in decision-making processes, while systems like BiasEye [38] and Lumos [47] emphasize the need for transparency and bias mitigation in evaluation practices.

The literature also highlights the importance of visualizing uncertainty, particularly when evaluating subjective data, such as personal statements and recommendation letters. Techniques from Dimara et al. [11] and Padilla et al. [52] provide insights into how decision-makers can better navigate cognitive biases and understand variability in their assessments. These strategies reinforce the need for transparent processes that account for the complexities of human judgment.

Overall, the research reviewed points to the necessity of integrating tools that balance weight-based decision-making and uncertainty visualization. This thesis builds on these insights to propose a framework aimed at enhancing the admissions process by incorporating comprehensive visualization techniques that ensure fairness, transparency, and alignment with academic objectives.

3 Pilot Study

A qualitative pilot study, using semi-structured interviews, was conducted to investigate the first sub-question of the main research question: How do admissions committees make final decisions, and what factors are considered when evaluating applications before making those decisions?

This pilot study interview involved interviews with nine experts from Utrecht University's admissions domain.

The goal was to gain a comprehensive understanding of the real-world practices of admissions professionals, specifically how they evaluate individual applications and assess them against established policies and criteria. The study aimed to uncover the thought processes, methods, tools, and challenges that these professionals encounter in their daily work. This understanding was crucial for developing visualization solutions that are closely aligned with the specific needs and workflows observed during this pilot study while acknowledging that these may vary across different institutions.

Reflecting on Munzner [45] *Nested Model* in visualization design [45], which underscores the importance of thoroughly understanding the domain problem before progressing to data abstraction and visualization design (see Figure 3), this pilot study played a pivotal role in informing the solution proposed in this thesis. The insights gathered during this phase not only shaped the development of EvaluationViz within the context of higher education admissions but also have broader implications for addressing challenges in multi-criteria decision-making across various evaluation contexts. By following the stages outlined in the Nested Model, the study ensured that the tool's design aligned with the needs of evaluators while maintaining a focus on clarity, transparency, and collaboration.

The following subsequent sections of this chapter detail the pilot study. Subsection 3.1 outlines the recruitment and participant selection process, while subsection 3.2 describes the materials and tools used during the interviews. The interview protocol and procedures are further elaborated in subsection 3.3. The key findings discussed in subsection 3.4 revealed that while final decisions are formally recorded in digital systems and documented through official submission forms, the processes leading to these decisions are often informal and lack structure. Many admissions professionals rely heavily on instinctual judgments, and crucial elements such as the strengths and weaknesses of an application, along with the inputs that influence final decisions, are not consistently transparent. Additionally, collaboration between evaluators often remains informal, typically limited to casual conversations and unstructured text-based communication rather than structured discussions when forming final decisions. This lack of a systematic decision-making process and reliance on informal collaboration underscores the need for a more structured approach that effectively captures and communicates evaluators' thought processes. Such an approach would ensure that the final decision-making process is well-supported, transparent, justified, and thoroughly documented.

3.1 Recruitment & Participant Selection

Participants for the expert interviews were recruited through purposive sampling [54], targeting individuals with specific knowledge and expertise in the admissions process within Utrecht University. Initially, potential participants were identified through professional networks and Utrecht University's official channels, focusing on those involved in handling admissions and policies, members of admissions committees, and course coordinators involved in evaluating applications. Participants were selected from the *Graduate School of Natural Sciences (GSNS)* [77] and the *Graduate School of Life Sciences (GSLS)* [76] department.

A recruitment email was sent to these individuals containing a link to a Qualtrics survey (see Attachment 1), which provided detailed information about the study's purpose, interview structure, and estimated duration. It also explained how their data would be used and protected. If they agreed to participate in the study and provided their consent, they were asked to submit their demographic details and contact information and select an available time slot for the interview. Snowball sampling [22] was also employed, where initial participants recommended other qualified individuals. The recruitment process continued until data saturation was reached—meaning that no new insights were emerging from the interviews [79]. Once participants expressed interest, they received a formal invitation, including their preferred interview times and additional study details, allowing them to ask any further questions or seek clarifications beforehand.

The final participant group included **9** individuals: 2 Policy Advisors, 1 Honors Program Coordinator, and 6 members from Admissions Committees and Program Coordinators. Of these participants, 7 were female, and 2 were men, with average ages ranging from 30 to 59. The participants brought a diverse range of experience, with their expertise in the admissions domain spanning from at least one and a half years to up to 10 years. Their professional backgrounds included fields such as Human-Computer Interaction (HCI), Artificial Intelligence (AI), Data and Computer Science, Game and Media, History and Philosophy of Science, and the Department of Biology.

3.2 Materials/Tools Used

Interviews were conducted either in person or online via Microsoft Teams, depending on participant preference. Sessions were recorded using both the Teams platform and a mobile device, with additional digital note-taking during the sessions. Although Teams offers transcription services, the quality was inconsistent, necessitating a thorough review of recordings to ensure note accuracy.

Graphs and tables developed from the analyzed data were created manually using Miro board workspace [43]. All raw and processed data were securely stored on Utrecht University's Yoda system, a research data management platform that ensures the security of research data, in compliance with the university's data management guidelines and data privacy.

3.3 Interview Protocol & Procedure

The semi-structured interviews were guided by a carefully designed protocol, detailed in 2, to ensure consistency and minimize external influences on data collection. This protocol was crucial for maintaining uniformity across interviews and facilitating rigorous and balanced data analysis.

The interview questions were designed to explore how admissions professionals evaluate each application, taking into account established policies, the complexities inherent in the evaluation process, and their methods of collaboration in making final decisions. The questions also aimed to assess their level of certainty in the process, the consistency of their approaches, the tools they use, the challenges they face, and any wishes they may have to further improve the process.

Inspired by a formative research approach similar to the BiasEye project [38], which also examines the admissions process and aims to make it more authentic (as discussed in the Literature Review), the interview questions were carefully compared and adjusted to eliminate bias and ensure comprehensive data coverage, ultimately informing the visualization solution proposed in this thesis.

The procedure began by welcoming the participants, introducing the project, and reiterating the data privacy protocols. To ensure consistency across all interviews, a standardized introduction script was used (see Appendix 2), guaranteeing that the same information was communicated clearly and accurately to all interviewees. The interview then proceeded naturally, with the interviewer (the author of this thesis / the study's researcher) guiding the discussion through the mentioned prepared questions as listed in Appendix 2.

3.4 Findings

The analysis of qualitative data from expert interviews was approached through an iterative and careful engagement with the raw data. This involved multiple readings of the transcriptions and notes, as well as re-listening to the audio recordings to ensure that no important details were missed. Consistent with the methodology of Straussian Grounded Theory [66], a systematic approach was employed to break down and interpret the data. The first step involved open coding [66], where interview responses were divided into meaningful segments and assigned descriptive codes. This method allowed for the identification of emerging concepts within the data, which were then grouped into broader categories, leading to the early identification of patterns and relationships. For example, some of the core challenges faced in the evaluation process were captured in the extracted quotes presented in Figure 12, which reflect the common concerns expressed by participants regarding the assessment of student competencies.

"[...] is the student good enough, can the student handle it, and do they know what's expected? [...]"
"[...] you will never be sure what the student is taught and at what level [...]"
"[...] no collaboration features or visualizations to assist in the decision-making process [...]"
"[...] not enough flexibility or automation in the system when dealing with international applications [...]"
"[...] language proficiency assessments don't fully reflect whether the student is truly competent in the language [...]"

Figure 12: Extracted quotes from participants highlighting challenges in the evaluation of student competencies

Following this, axial coding was used to explore the connections between these categories [66]. This phase involved mapping out how different codes related to one another, facilitating the development of central themes that offered more cohesive insights. During this stage, the detailed process descriptions and the time spent on each activity by the participants (as illustrated in Figure 13) helped to inform the construction of the overall process flow. This flow was later visualized in the process flowchart in Figure 14, which provides a comprehensive view of how decisions were made throughout the evaluation process and common data used.

Step	Process Description	Time Spent per Applicant
Opening Osiris and Initial Review	*"Open **Osiris** and manually review applications, checking documents **one by one**"*	*"**5-10 minutes** for local applicants, **30 minutes** or more for international applicants"*
Reviewing Transcripts and Grades	*"The focus is primarily on applicant data such as **grades and transcripts** for the **required courses**, along with **CVs** and **motivation letters** to understand the applicant's intent. In some cases, participants may request **references** or **additional writing samples** for further clarity"*	*"5-10 minutes per local applicant, with international applications taking significantly longer"*
Checking Additional Documents	*"Log into Osiris, **download attachments**, and review transcripts and motivation letters one at a time"*	*"Applications are processed one at a time; decision times vary depending on the complexity of each case"*
Collaborative Review	*"After downloading files from Osiris, participants often **share them with colleagues for collaborative decision-making**. They noted that not all team members can view files at the same time, which can extend the review process. One individual typically finalizes the decision, but group input is especially important when dealing with unfamiliar academic backgrounds, such as multidisciplinary cases"*	*"Collaborative reviews can take up to 30 minutes or, in some instances, stretch over several days."*
Verifying Student's Academic Background	*"Participants evaluate courses based on their titles. If the applicant's institution or course is unfamiliar, they turn to the **institution's website** or other **public sources for additional information**"*	*"Processing takes longer for **interdisciplinary** candidates, as the evaluation process is more complex"*
Finalizing the Decision	*"The process usually starts with a review of the applicant's **previous university reputation**, followed by a detailed examination of CVs, transcripts, and **course content**. Complex cases require more time than straightforward ones"*	*"For most applications, this takes 10-15 minutes, though more complex cases require additional time"*

Figure 13: Process description table and time spent on various tasks during the evaluation, as reported by participants

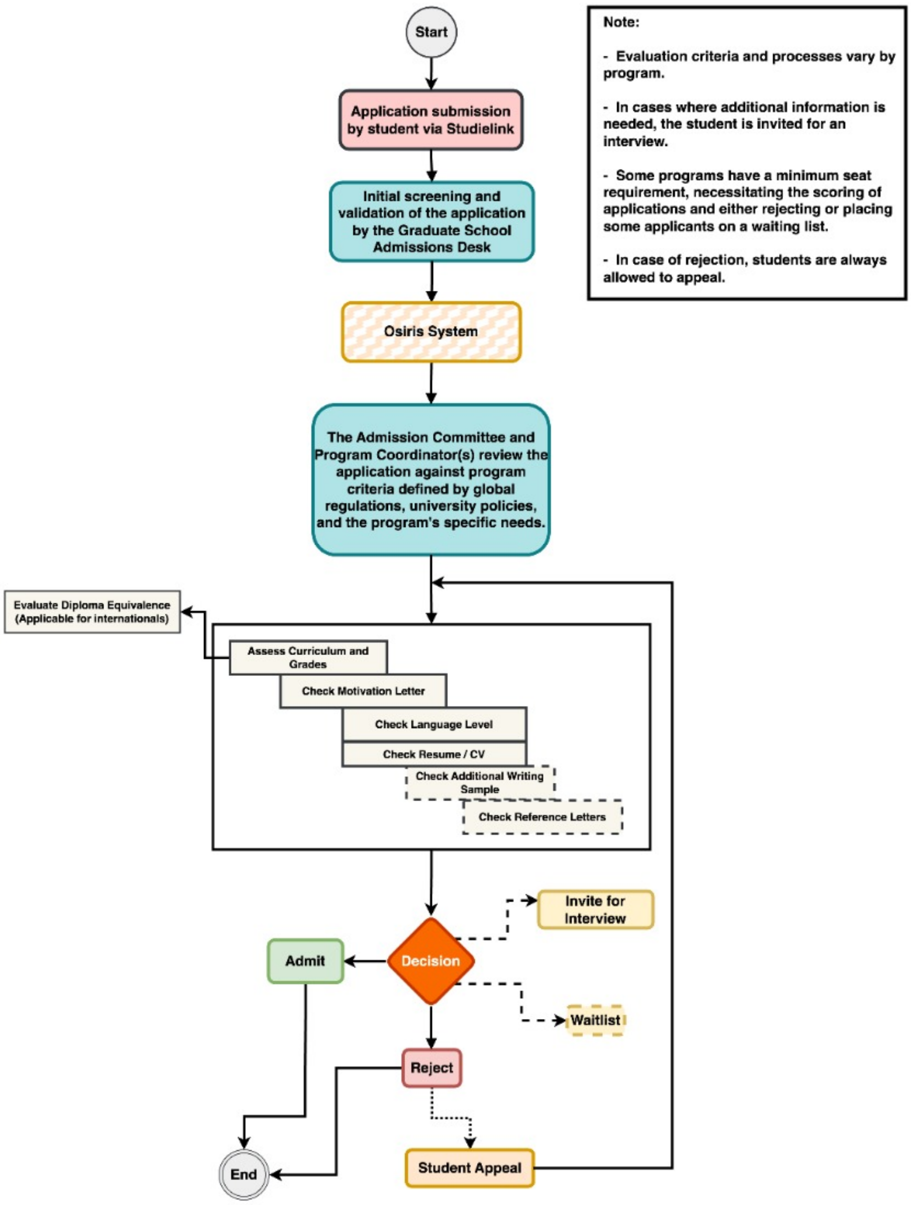


Figure 14: Process flowchart illustrating the overall decision-making process data used during student evaluation

The inclusion of two persona descriptions (outlined in 15) served to illustrate the different roles, goals, challenges, needs, and user scenarios within the evaluation process. One persona represented a more complex scenario, while the other depicted a simpler flow. These personas helped to highlight how decision-making and task distribution may vary across different contexts and provided insight into specific areas where improvements could be made, such as enhancing collaboration and addressing workflow inefficiencies.

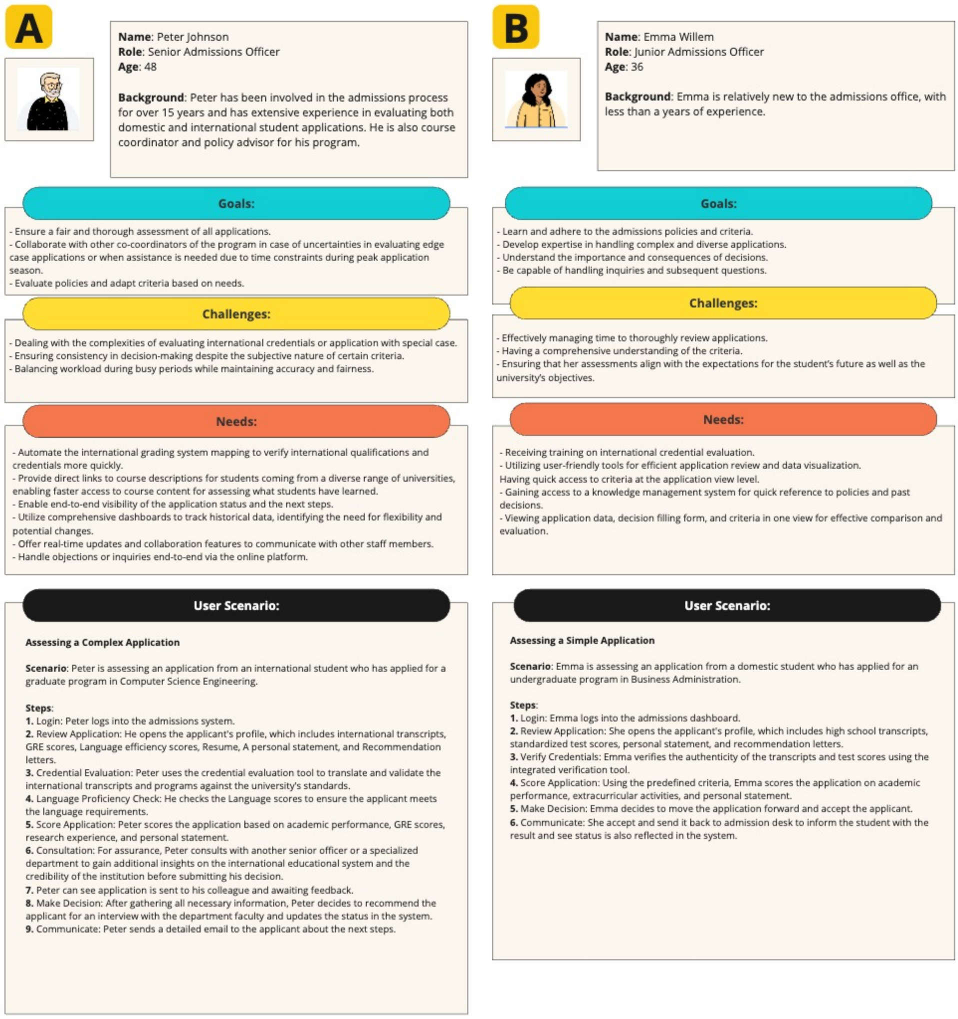
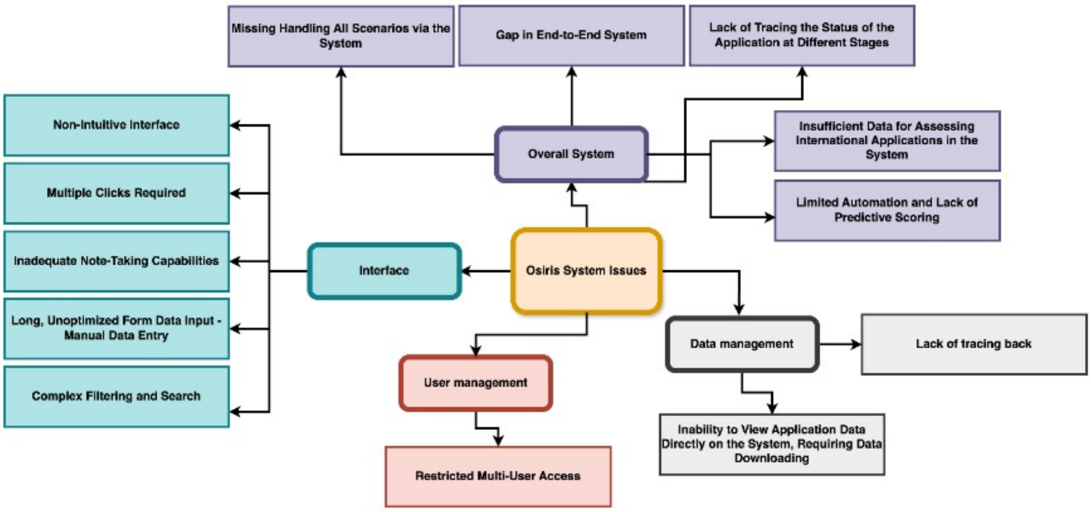


Figure 15: Persona illustrating roles, goals, challenges, and needs in (A) a complex scenario and (B) a simple scenario

Finally, selective coding was conducted to refine the emergent categories, focusing on identifying core themes that captured the essence of the participant's responses [66]. This stage also involved categorizing the issues highlighted by participants regarding both the tools and the process. These issues, grouped by category and linked to their respective impacts, are summarized in Figure 16. Additionally, the *'magic wand'* suggestions from participants, which represent their wishes for changes or improvements in the system, are captured in Figure 17. These suggestions reflect a desire for more intuitive tools and better support for decision-making, aligning with the broader themes identified through the analysis.



Issue Category	Issue Description	Impact
Flexibility	*"Osiris lacks flexibility in processing diverse international education systems"*	*"Hinders the evaluation of international candidates"*
Missing Data	*"International transcripts often lack data for proper evaluation"*	*"Delays decision-making and requires more manual work for evaluators"*
User Interface	*"Non-intuitive interface, multiple clicks required, and complex search filtering"*	*"Slows down processing and increases user frustration"*
Data Management	*"Limited automation and lack of predictive scoring for applications"*	*"Results in slower decision-making and requires more manual work"*
Collaboration	*"Multiple reviewers cannot collaborate easily in Osiris"*	*"Slows down decision-making and increases workload"*
Note-Taking Capabilities	*"Inadequate note-taking capabilities and long, unoptimized form input"*	*"Adds unnecessary workload and reduces system efficiency"*
User Management	*"Restricted multi-user access and lack of tracing across application stages"*	*"Reduces efficiency in managing and reviewing applications"*
Transparency	*"Lack of clarity in admission processes and what criteria are most important"*	*"Difficulties for both students and administrators to understand rejections"*
Technology Overload	*"Excessive information in Osiris, but it's not fully utilized"*	*"Leads to cognitive overload, making decision-making harder"*
AI Usage Concerns	*"Concerns about AI-generated writing samples affecting fair assessments"*	*"Raises doubts about authenticity in student applications"*

Figure 16: Issues mindmap and table identified by participants regarding the evaluation process and tools, grouped by category and their impacts

If you had a 'magic wand' to create a tool or implement a change in the admissions process, what would it be? How would you envision this tool assisting you and your team?
*"I wish there were tools for multi-user collaboration in Osiris to enable smoother **collaborative decision-making** and improve **data sharing**."*
*"I wish there were better systems for **status and decision tracking**, so we could easily monitor and manage application progress"*
*"If I had a magic wand, I would create a standardized **global education data** table for comparison, including a global statistics database to show where students come from and enable easy **data comparison**"*
*"I wish we could improve **language proficiency assessment tools** to provide more accurate and reliable metrics for evaluating applicants"*
*"If I had a magic wand, I would **automate the generation of rejection letters** to save time when processing rejected applicants"*
*"I wish we could introduce **AI and automation** to assist with evaluating international applications and **mapping** them more efficiently"*
*"If I had a magic wand, I would standardize admission criteria across programs and use **data-driven approaches** like signal detection theory to improve decision-making"*
*"I wish we could improve **transparency** in the admissions process and make it easier to **filter and search** through student applications"*
*"If I had a magic wand, I would streamline **communication bottlenecks** and reduce the need for manual emailing, especially during peak periods"*

Figure 17: Participant suggestions on improvements they would like to see in the system, based on the 'magic wand' exercise

It is important to note that each program prioritized different data points (Grade transcripts, CV, etc.) or aspects of applicants, so a combination of all relevant general data points was selected for implementation in the proposal offered by this thesis for visualization solution. This approach accommodated the variations in criteria across different programs, ensuring the tool's flexibility and utility in multiple contexts.

The findings from the pilot study revealed several key challenges faced by admissions professionals, particularly regarding transparency, consistency, and the efficiency of the decision-making process. A repeated issue raised by participants was the absence of structured tools to support both individual and collaborative evaluations. As one admission expert commented, "There are no features to assist in the decision-making process," leading to informal and undocumented discussions. This lack of formal mechanisms not only complicated the evaluation process but also made it difficult to ensure fairness and accountability in decision-making.

Furthermore, some admissions officers raised concerns about the evaluation of international applicants, highlighting the limitations of existing systems such as Osiris, which struggled to accommodate unfamiliar academic backgrounds and non-traditional qualifications. The inflexibility of such systems, combined with the need for evaluators to rely on instinct when faced with ambiguous cases, underscored the need for more data-driven tools to streamline the evaluation process and reduce the reliance on subjective judgments.

In addition, a deeper analysis of the data revealed that a significant amount of time was spent searching for external information. Many participants described having to consult institutional websites or other public sources to evaluate applicants whose academic backgrounds were unfamiliar to them. They also mentioned that there is always an element of uncertainty regarding whether the student is a good fit, what they have learned before, or if they know what to expect. This added layer of research not only complicated the workflow but also exposed a gap in the tools available for efficiently accessing and verifying external data during the admissions process. As one participant explained, "Processing takes longer for international and interdisciplinary candidates, as the evaluation process is more complex." This reliance on external information increased the cognitive load on evaluators and further slowed the decision-making process. An analysis of the time spent per applicant revealed a noticeable difference between the time spent evaluating familiar domestic applicants and that spent on international and interdisciplinary applicants, where additional research was often required.

The need for improved collaboration was another prominent theme. Although the evaluation process involved multiple stakeholders, collaboration between evaluators remained informal and inconsistent. Participants expressed a desire for more structured and transparent tools that could facilitate group discussions and decision-making. One participant remarked, "There are no collaboration features or visualizations to assist in the decision-making process," highlighting the gap in tools that support communication and joint decision-making. This lack of standardized procedures meant that evaluators frequently defaulted to intuition or personal experience, rather than following a uniform process. This contributed to variability in the assessments of similar applicants, further emphasizing the need for a more formalized, structured decision-making framework.

Moreover, participants identified challenges in evaluating qualitative aspects of applications, such as personal statements and recommendation letters. The subjective nature of these documents made it difficult to apply consistent evaluation criteria, leading to varied interpretations across evaluators. In cases where applicant qualifications were ambiguous or unclear, evaluators often relied on informal judgment rather than following a transparent and standardized process. This underscores the need for tools that can standardize the evaluation of qualitative data and provide clear criteria for decision-making, particularly in complex or borderline cases.

The gaps identified throughout the admissions process—whether in collaboration, transparency, or the evaluation of external and qualitative data—highlight the urgent need for tools that can support evaluators in making holistic and well-documented decisions. The informal and fragmented nature of the current evaluation process, as described by participants, underscores the importance of developing solutions that integrate all aspects of evaluation—from external data sources to structured group decision-making. Addressing these gaps holistically would ensure that decisions are more transparent, better supported by data, and consistently applied across applicants.

The findings from this pilot study, particularly in the areas of structured holistic evaluation, collaboration, and the application of information visualization principles, directly influenced the formulation of the research question and the development of the visualizations in this thesis. This visualization solution was designed to address some of the challenges shared by experts and incorporate insights from the literature review. These inputs were incorporated into the proposal to create a more structured and transparent decision-making process.

The next chapter 4 introduces the EvaluationViz and provides detailed information on how these insights shaped the proposed solution in the thesis, discussing the design, functionality, and justification of the choices.

4 EvaluationViz

EvaluationViz is a custom one-page evaluation form and visualization developed by the author to support the assessment process in university admissions. It enables evaluators to systematically review critical data points about applicants and synthesize individual assessments into a comprehensive profile. The tool focuses on the intermediate stages of decision-making, specifically on evaluating individual components of applications, while excluding other aspects of the end-to-end administrative workflow. EvaluationViz is built to address the research question: *How can data visualization support intermediate decision-making in university admissions, particularly in evaluating individual aspects of applicants and facilitating collaborative final assessments among evaluators?*

EvaluationViz features a form that contains application data and collects three metric inputs—score, weight, and uncertainty—for each data point. The ***score*** represents the overall perceived rating for each application data point, ***weight*** reflects the perceived significance of each data point in the decision process, and ***uncertainty*** indicates the level of confidence in the assessment. The choice of these metrics, along with their justification, is further elaborated in subsection 4.2.

The visualization section includes four key visualizations, each offering a unique perspective on the collected input. These visualizations range from summarizing data into a composite overall application score, displaying a ***Radar/Spider Summary Chart***, and a ***Tabular Panel Bar Chart*** that presents composite values or filtered data points based on specific metrics for each application component (hereafter referred to as the ***Composite Score***, with the formula detailed in subsection 4.2). Another visualization, a ***Stacked Bar Chart***, illustrates how each component contributes to the overall score per applicant for each advisor, allowing a clear comparison of the weighted contributions. Finally, the ***Lollipop Chart*** provides a detailed breakdown of each advisor's input across the three metrics—score, weight, and uncertainty—based on filtered application data points. The design, interactions, justifications, and alternative design considerations for all visualizations are thoroughly discussed in subsection 4.2.

The design and framework of EvaluationViz were informed by literature on information visualization and insights from an initial expert pilot study. Additional feedback was obtained through a structured session involving seven participants selected through convenience sampling [17]—six men and one female (average ages between 20 to 49) with backgrounds in software engineering, business management, and UX design, representing German, Dutch, Colombian, American, and Indian nationalities. Two evaluators had prior knowledge of human-computer interaction (HCI) concepts, while six were familiar with information visualization principles. This feedback session was vital to validate the clarity of the design before the actual tool evaluation addressing the research question. Prior to the session, participants signed consent forms outlining the project's purpose and data privacy protocols, alongside the structured questionnaire described in 3. Each evaluator interacted with the tool on the author's laptop, entered data, and provided feedback via the questionnaire. Additionally, participants completed a demographics survey, as outlined in 6.

The session aimed to evaluate the clarity, effectiveness, and utility of the tool's visualizations and overall design. Participants interacted with several visualizations, including an overview chart, initially presented as a lollipop group bar chart, which was later redesigned into a radar chart based on feedback from 5 out of 7 participants who found the radar chart more effective in showcasing application strengths and weaknesses.

Other visualizations, such as a tabular panel bar chart (with interactions improved post-feedback), stacked bar charts, and lollipop charts, were also explored. Adjustments to these visualizations, discussed in the subsection 4.2, detail their iterative development. The visualizations emphasized different aspects of the applicant data and facilitated collaborative decision-making among evaluators. Feedback from this session informed further refinements to EvaluationViz, enhancing usability, transparency, and the tool's ability to provide data-driven insights for final evaluations.

Subsection 4.1 will explore the technologies used to build EvaluationViz, while the design choices, evaluation form, and visualizations are discussed in subsection 4.2.

4.1 Technology Used

EvaluationViz was built using the **D3.js** library, which allowed for the creation of dynamic and custom-made from scratch visualizations tailored to the specific requirements of the project.

Next.js, a React-based framework, was used for the application's architecture, offering efficient state management and handling of real-time user input. This ensured that evaluators could interact with the tool in a responsive manner during the evaluation process.

Tailwind CSS, combined with Daisy UI, was employed for visual styling and the user interface. Tailwind's utility-first design approach allowed for streamlined custom styling, while Daisy UI's pre-built components helped simplify the development of the evaluation form.

4.2 Design

4.2.1 Data

The dataset utilized in EvaluationViz revolves around **seven key application components** that are widely considered in the university admissions process: **Previous Education, Grades Transcript, Language Proficiency, CV, Motivation Letter, Reference Letters, and Writing Sample**. These data points were selected to reflect common and important criteria in admissions decisions, providing a relevant and useful evaluation platform.

Due to the lack of access to actual admissions data, **two simulated applicants** were created, each with anonymized but realistic data spanning the aforementioned categories (see Appendix 7). This simulation ensured that EvaluationViz could effectively mimic a real-world admissions scenario, providing evaluators with meaningful content to assess without compromising privacy or ethical standards.

In order to evaluate the simulated applicants, **three advisor profiles** were also generated. These profiles represented different evaluators within the system, each tasked with providing feedback across the various data points.

The evaluators were then asked to assess the applicants using three critical metrics: **Score, Weight, and Uncertainty**, each ranging from **0 to 5**. Score reflects the perceived quality of the applicant's data, Weight indicates the importance of the data point in the overall decision-making process, and Uncertainty highlights the evaluator's confidence in their judgment. To ensure independent evaluations, the data was reset between assessments, allowing for an isolated analysis of each advisor's input. This choice is supported by research in information visualization, which emphasizes the importance of clarity and simplicity in decision-making processes. Studies have shown that introducing more metrics can lead to information overload, complicating the evaluation process without necessarily improving decision accuracy or adding significant value to the final outcome [10].

The **simulated rubrics policy criteria** given to participants helped simulate the role of an admissions officer, guiding their evaluations while allowing room for subjective interpretation (see Appendix 7).

Moreover, in addition to visualizing the input data through various views tailored for different tasks, the overall score of each application data point (application component such as CV) is calculated using a **Composite Formula** that integrates the score, weight, and uncertainty, as illustrated in Figure 18. The formula,

$$s \times w \times \left(1 - \frac{u}{5}\right)$$

, combines these elements in a multiplicative manner to reflect their interdependent roles in decision-making. By multiplying the score (s) by the weight (w) of each criterion, the formula ensures that more important factors have a greater influence on the final result, as supported by prior research emphasizing the importance of dynamic weight adjustments in multi-criteria decision-making tools like WeightLifter and LineUp [53, 23]. The inclusion of uncertainty (u) in the form of

$$\left(1 - \frac{u}{5}\right)$$

ensures that higher uncertainty appropriately reduces the composite score, mitigating the risks associated with overconfidence in uncertain data, a principle underscored by Dimara et al. [10] and Padilla et al. [52]. This approach is grounded in the literature on cognitive bias in visualization, which emphasizes the importance of representing uncertainty to prevent misinterpretations and cognitive distortions in decision-making [11]. By scaling the score based on uncertainty, the formula provides a balanced reflection of both the evaluative data and the level of confidence in that data, ensuring more reliable and transparent decision outcomes in university admissions.

Figure 18: Composite Score Formula, also displayed in the frontend when data is filtered by the composite score rather than specific metrics, to enhance transparency.

4.2.2 Choice of Colors, Marks & Channels

The design choices in EvaluationViz, particularly in the use of color, marks, and visual channels, were based on principles of data visualization.

One of the critical considerations was color assignment, which was vital for distinguishing different data categories visually. Based on research in categorical data visualization, it was important to assign a consistent color to each of the seven data points across all visualizations, ensuring that users could easily correlate information between the evaluation form and the visual elements [37]. The use of D3.js provided precise control over these color assignments, ensuring that the differences between adjacent data categories were perceptually clear and harmonious, as discussed in Li et al. [37]'s optimization approach for color distinguishability in visual charts. After several iterations of the color scheme, the final selection—shown in Figure 19—was chosen based on feedback from the evaluation session, which favored a more matte color palette.

Optimizing color assignment was particularly important in charts containing adjacent blocks, as it improves readability and enhances the user experience. Ensuring that colors were distinguishable even when adjacent, allowed evaluators to easily interpret the information without confusion or fatigue [37].

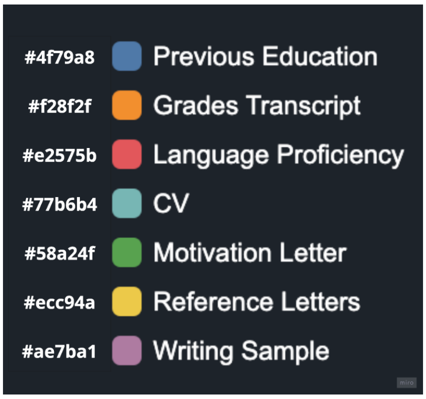


Figure 19: Final color scheme selected after multiple iterations, incorporating feedback from the evaluation session that favored a matte color palette for better visual clarity.

Moreover, dark mode was adopted as the primary visual theme for EvaluationViz to reduce eye strain and enhance contrast. Research suggests that dark mode can be more comfortable for prolonged viewing, particularly in environments where visual fatigue may be a concern [81].

Additionally, different colors were selected for the metrics—score, weight, and uncertainty—which were distinguishable on top of the seven categorical colors for the application data. This ensures that the three metrics can be quickly identified without blending with the other categories. See Figure 20 for the final metric color scheme.

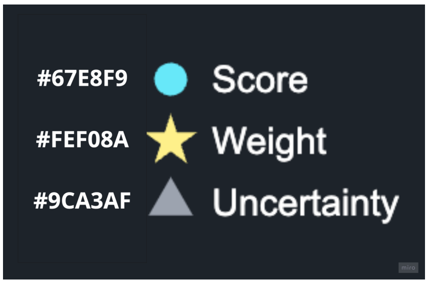
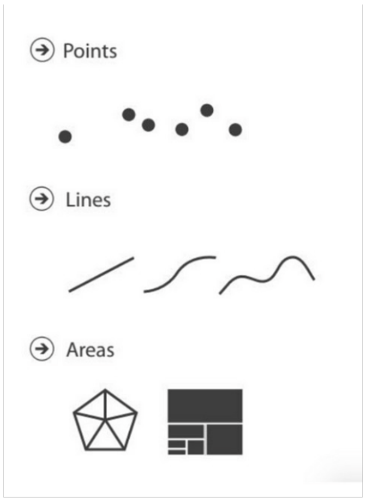


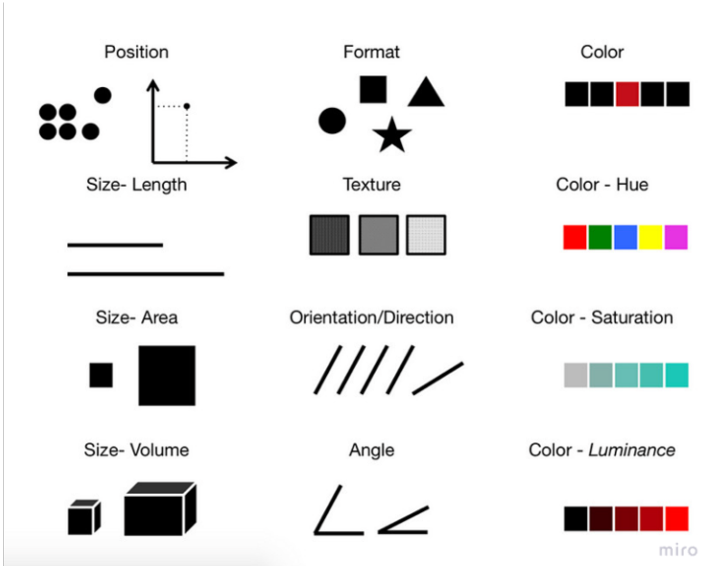
Figure 20: Final color scheme for the metrics, chosen to ensure distinction from the seven application data categories. The visual marks used are shapes (circle for Score, star for Weight, and triangle for Uncertainty). The channels employed are color hue to differentiate the three metrics, with #67E8F9 (blue) for Score, #FEF08A (yellow) for Weight, and #9CA3AF (gray) for Uncertainty. The shapes provide an additional shape channel to aid recognition.

It is worth mentioning the need to avoid the overuse of colors and to separate visualizations from UI elements, where the color white was specifically chosen for all UI-related components.

In addition to color, specific marks—such as circles for scores, stars for weights, and rectangles for uncertainties—were used to visually encode these metrics. These design decisions were guided by Munzner [46] framework on visual channels, ensuring that EvaluationViz was not only visually appealing but also optimized for cognitive processing. The use of ordered attributes, such as length, facilitated clear comparisons of quantitative data, while categorical distinctions were enhanced through identity channels like color and shape, drawing inspiration from Figures 21b and 21a in Munzner's framework. Figure 22 also served as a reference for the ranking of visual channels according to their data type.



(a) Marks



(b) Channels

Figure 21: Marks and Visual Channels extracted from Munzner [46], illustrating the various methods for encoding data in visualizations.

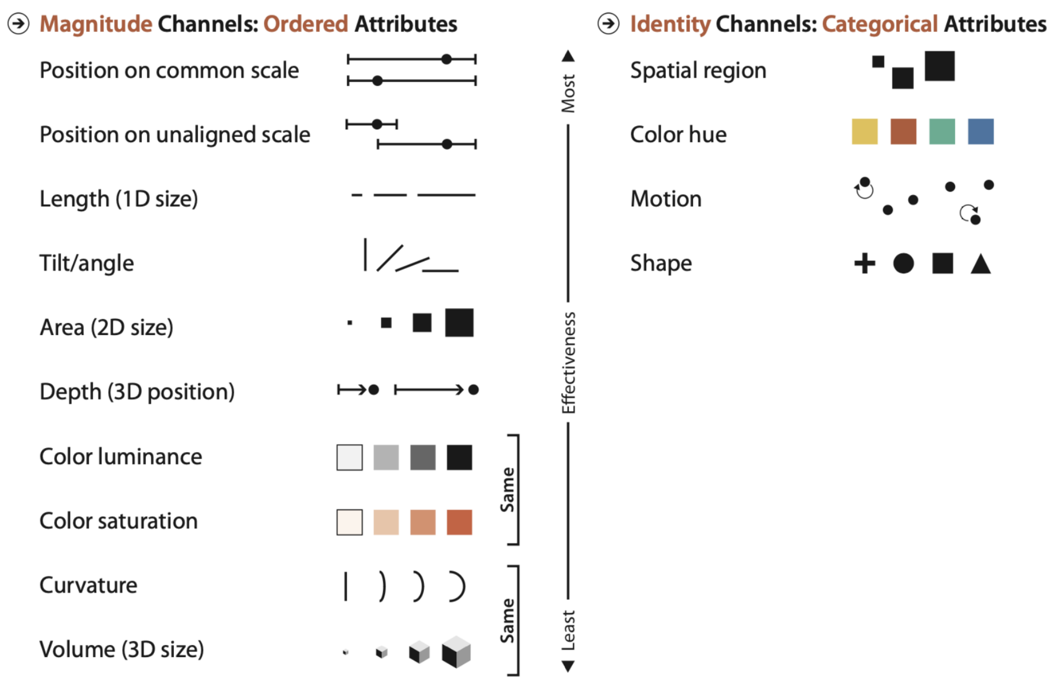


Figure 22: Ranking of channels according to their data type extracted form Munzner [46].

4.2.3 Evaluation Form Section with Range Inputs

The Evaluation Form Section in EvaluationViz is designed to capture an evaluator's judgment on key applicant data points, using interactive range sliders to input values for Score, Weight, and Uncertainty, and to collect raw data for the visualization section (see Figure 23 and 24). The form ensures a structured evaluation process by asking the evaluator to rate each application criterion—such as Previous Education, Grades Transcript, Language Proficiency, CV, Motivation Letter, Reference Letters, and Writing Sample—on a scale from 0 to 5. The Score reflects the perceived quality or relevance of the data point, the Weight captures its importance relative to other criteria, and the Uncertainty indicates the evaluator's confidence in their judgment. Each application data component includes a button that opens a modal showing detailed data, along with the policy and criteria for the program (see Figure 25).

Figure 23: Full view of the evaluation form consisting of seven sections, each representing an application data point. Inside each box, three range inputs are provided for score, weight, and uncertainty, with a button to display a modal showing application details and the corresponding rubric criteria. Additionally, a comment section allows advisors to provide further insights.

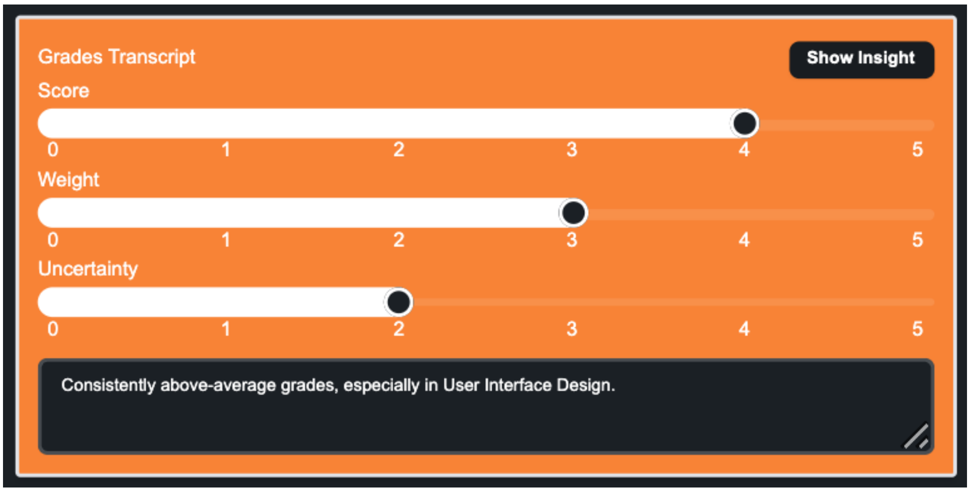


Figure 24: Zoomed-in view of one section (Grader Transcript) from the evaluation form for a clearer visual representation.

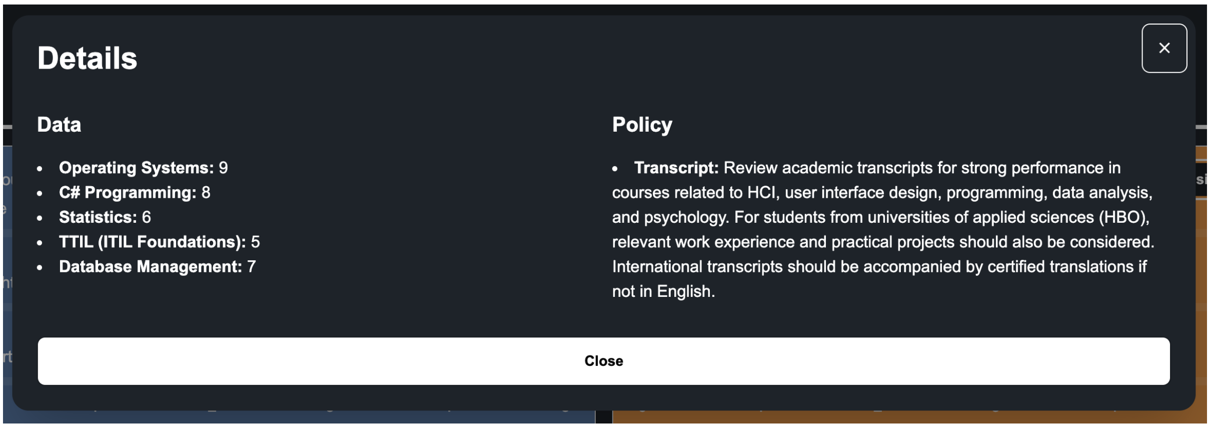


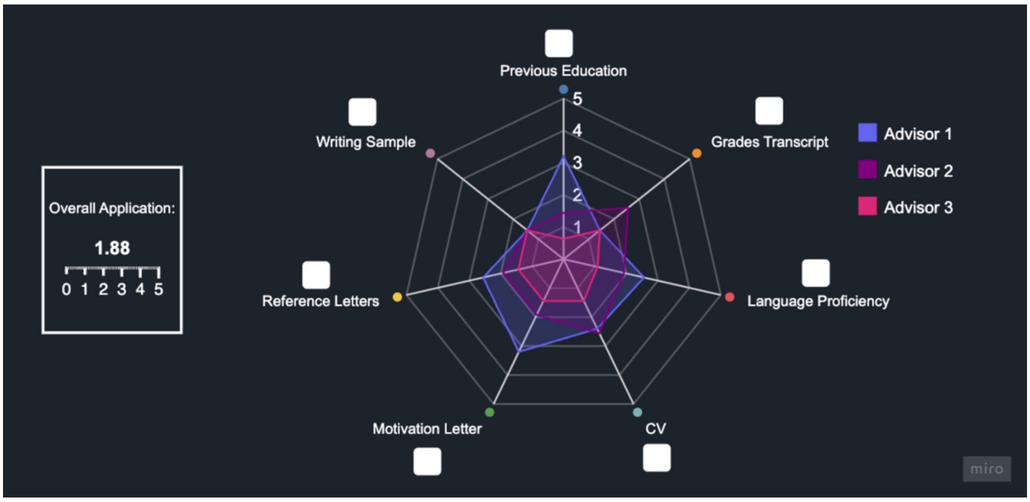
Figure 25: Modal displaying insights with application data and corresponding rubric details.

4.2.4 Visualization Section

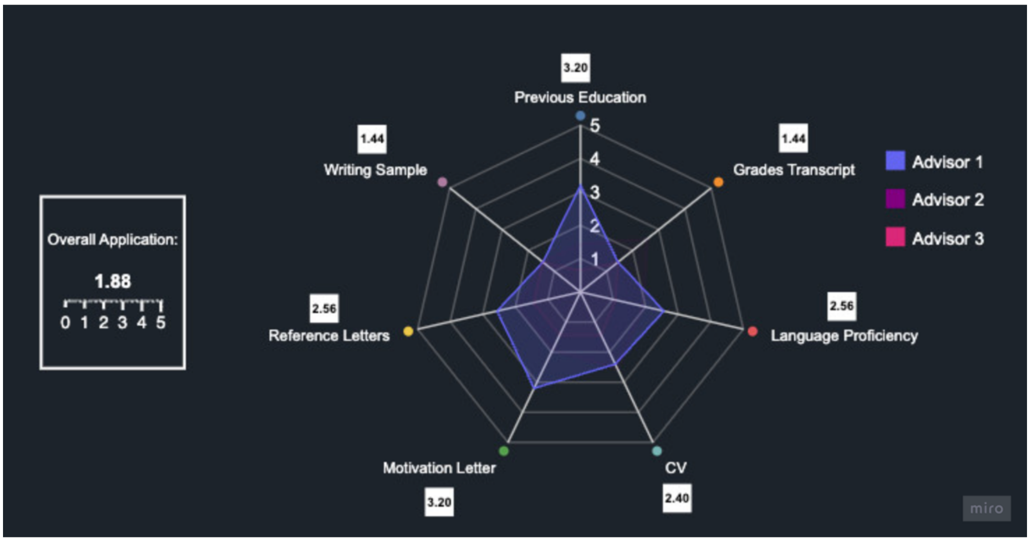
The Visualization Section of EvaluationViz includes several visualization types, each designed to clearly and effectively display the raw input data collected from evaluators for each application component, facilitating collaborative decision-making. The four main visualizations are the Radar/Spider Summary Chart, Tabular Panel Bar Chart, Stacked Bar Chart, and Lollipop Bar Chart, each serving distinct purposes in the evaluation process.

Radar/Spider Summary Chart:
The Summary Chart, implemented as a radar/spider chart, provides a holistic view of the applicant's evaluation across all seven data points, allowing evaluators to compare how each advisor rated individual criteria. By plotting each data point along an axis and connecting them, the radar chart visually highlights strengths and weaknesses from each advisor's perspective. Distinct colors are used for each advisor, ensuring that their assessments can be compared directly. The radar chart's ability to present multi-dimensional data in a concise form is one of the key reasons it was selected, as it excels at showing contrasts between advisors' evaluations at a glance—this need was repeatedly emphasized in the

visualization feedback study. The calculation is based on the composite score of each application component per advisor, with the chart displaying both individual component scores and an overall score that combines all component scores into a single value for each advisor. Finally, all advisors' evaluations are aggregated into one final overall score for the applicant, ranging from 0 to 5, to reflect the overall rank.



(a) Radar/spider chart displaying the overall evaluation of an applicant across seven data points based on composite input from three advisors.



(b) The same radar/spider chart shown during hover action, highlighting specific data points for detailed comparison, with data shown in white boxes on hover.

Figure 26: Radar/spider chart of overall collected input of applicant evaluations, comparing the static view (a) with the interactive hover action (b). The visual marks used are lines (connecting the evaluation points on each axis) and filled areas (shaded regions representing the overall evaluation for each advisor). The channel of position is used to represent scores along each axis, where each axis corresponds to a distinct application data point. Circular markers maintain the same color consistency across the entire EvaluationViz and the evaluation form. Color hue is used to differentiate between the three advisors, with distinct purple, blue, and pink shades.

Before implementing the radar/spider chart as the primary visualization tool for summarizing applicant evaluations, an alternative chart design, as shown in Figure ??, was initially considered. This design utilized marks and channels to represent key metrics, with distinct visual encodings employed to display the various dimensions of each advisor's evaluation. An arrow direction was used to indicate how uncertainty negatively impacted the score, while weight added extra value to the overall composite score. Additionally, a series of checkboxes allowed users to filter out specific application components (such as Motivation Letters or CVs) and observe how these adjustments influenced both the overall composite score and individual advisor assessments. The ability to filter application data provided valuable insights into how different components contributed to the evaluation. This flexibility was well-received during feedback sessions, though it was seen as more suitable for evaluating policy decisions or analyzing the impact of application data. Feedback from the same sessions indicated that the radar/spider chart was more effective at communicating the strengths and weaknesses of applicants at a glance. The radar chart's ability to map multiple dimensions onto a unified, easy-to-interpret visual was more aligned with user expectations for assessing overall applicant profiles, making it the preferred choice. Therefore, while the alternative design offered detailed analysis, it was ultimately replaced by the radar/spider chart, which better met the need for summarizing overall applicant evaluations concisely and intuitively.

It is worth noting that the initial design used the color red to represent uncertainty. However, research on color perception in uncertainty visualization has shown that neutral tones, such as grey, are more effective in representing uncertainty without implying an error or high-stakes situation, as red typically conveys [52]. This insight led to changing the uncertainty representation from red to a grey-shaded area, providing a more intuitive and less biased visual cue.

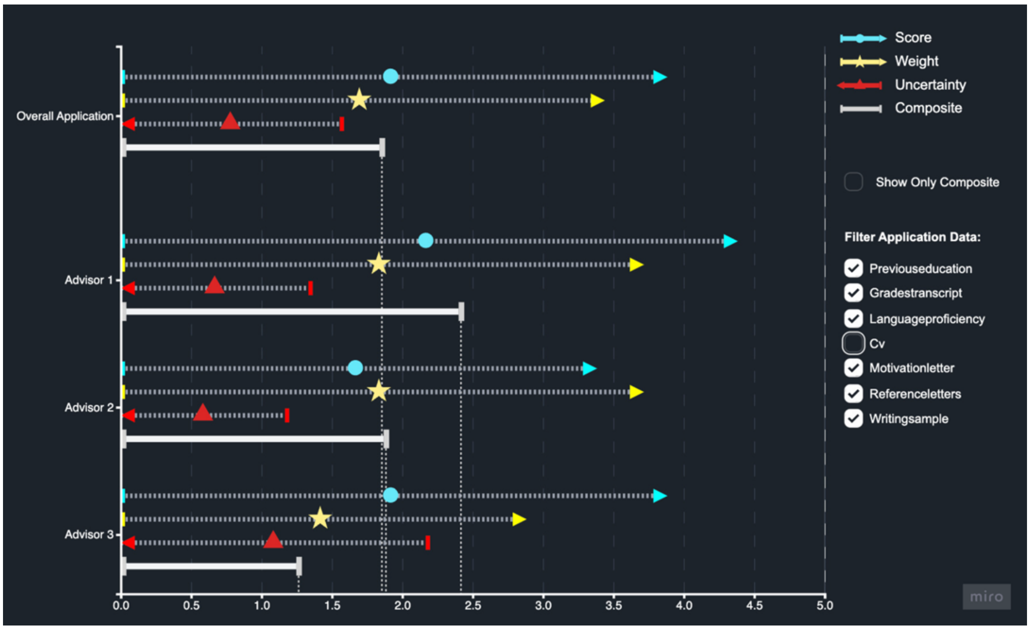


Figure 27: Static version of the alternative summary chart, which was initially explored before feedback sessions led to the radar chart design. The chart encodes score, weight, and uncertainty using distinct visual marks (circle for score, star for weight, and triangle for uncertainty). Dotted lines separate individual metrics from the composite score, while arrows indicate positive or negative effects on the score. The x-axis shows values from 0 to 5, while the y-axis lists overall applicant scores and individual advisors.

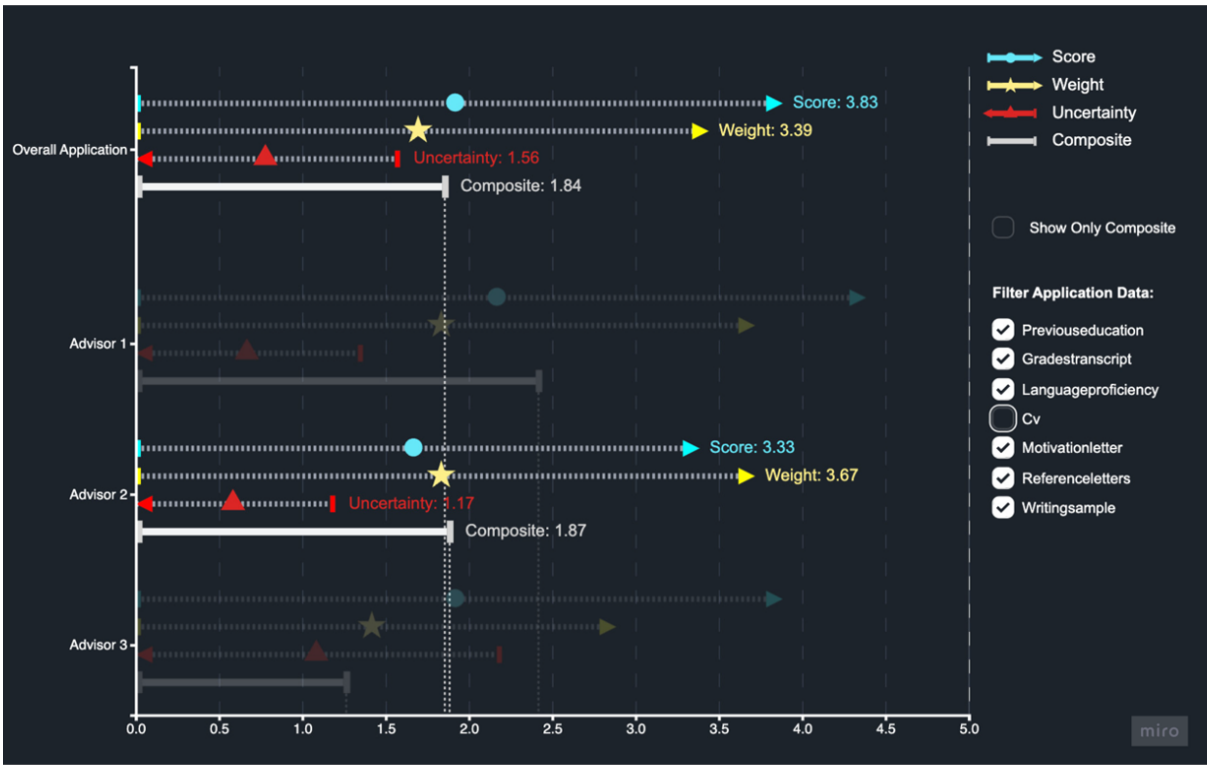
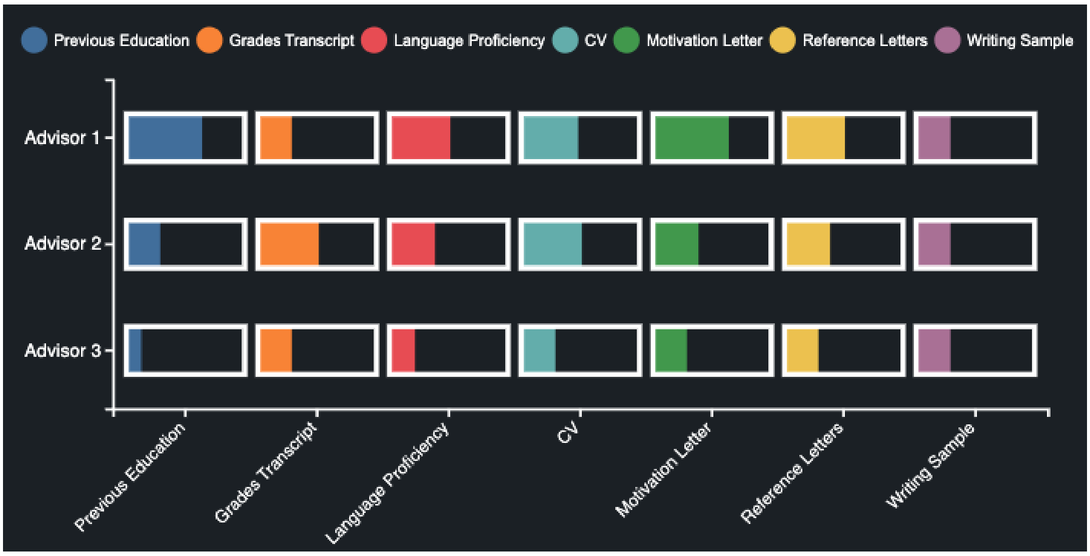
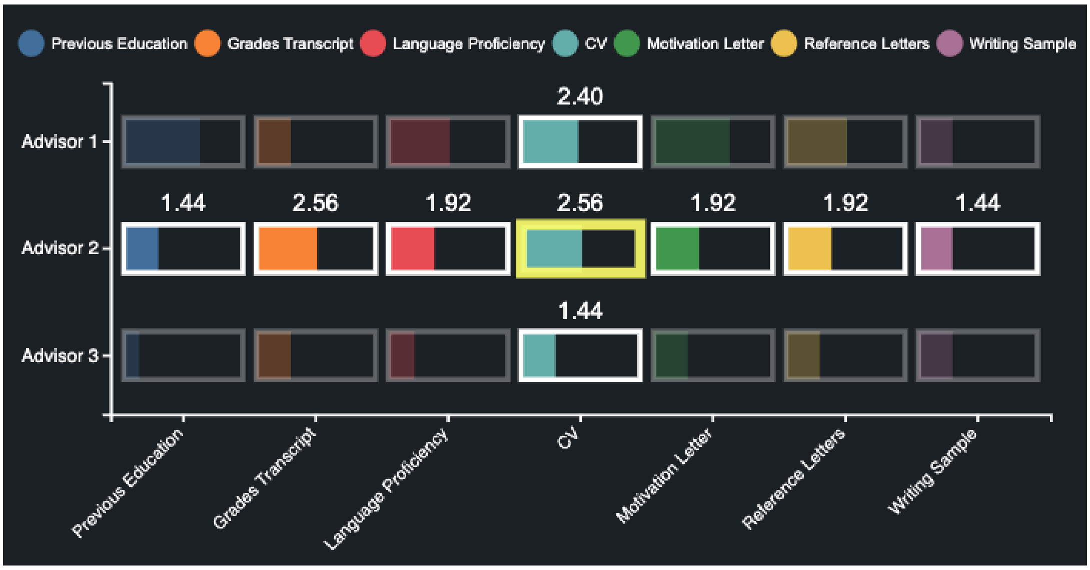


Figure 28: Hover effect displaying additional details, with values labeled along each line. Labels appear during hover to clarify metric values for score, weight, and uncertainty. The filter checkboxes allow users to highlight only the composite score and selectively remove individual application data to observe their impact on the overall score.

Tabular Panel Bar Chart:
The Tabular Panel Bar Chart (see Figure 29) breaks down the evaluation data into individual bars for each data point across different advisors. Depending on the selected filters from the dropdown menu, it displays either specific metrics per advisor for each application component or the composite score. This chart enables quick identification of which advisor has provided significantly higher or lower ratings for a given data point, allowing for easy detection of outliers or inconsistencies between advisors. All the boxes are of equal width and height, with the value represented by the filled portion of the box, ranging from 0 to 5. This design leverages the perception of length and height to emphasize value differences [46].



(a) Static view of the Tabular Panel Bar Chart, providing an overview of advisor evaluations for each application component. The visual marks used are rectangular bars, each representing an evaluation per application data. The x-axis represents the different application components (e.g., Previous Education, Grades Transcript, etc.), while the y-axis lists individual advisors. The same consistent channel of color hue is used to encode different application components, with each box color-coded to represent a specific component. The length of the filled portion of each bar represents the score (ranging from 0 to 5) given by each advisor.



(b) Hover effect view of the Tabular Panel Bar Chart, showing additional details such as the exact scores. When hovering, a cross-check function highlights how a specific component was evaluated across all advisors (horizontally) and how a particular advisor rated all components (vertically). This aids in identifying discrepancies and outliers.

Figure 29: Tabular Panel Bar Chart comparing the static view (a) with the interactive hover effect (b). The static view presents an overview of advisor evaluations for each application component, using rectangular bars as marks with color hue and length as visual channels. The hover effect provides exact values and cross-check functionality, allowing evaluators to compare scores across advisors and components.

The initial hover effect of the Tabular Panel Bar Chart included the addition of comparison lines to enhance visual clarity, especially since the data points were limited (see Figure 30); however, after feedback sessions, this hover effect was rolled back, and the chart reverted to its original form, as seen in the previous Figure 29.

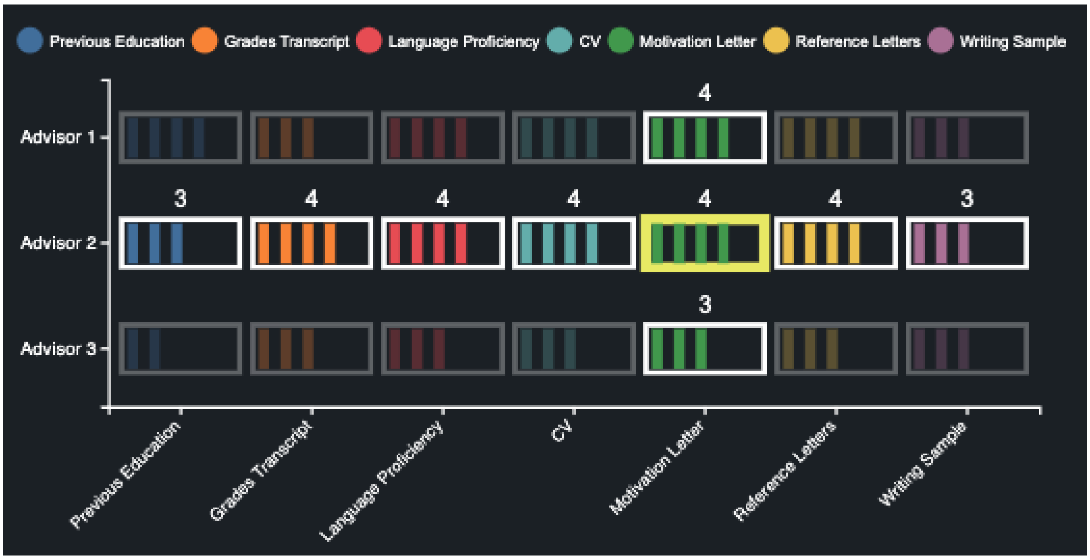
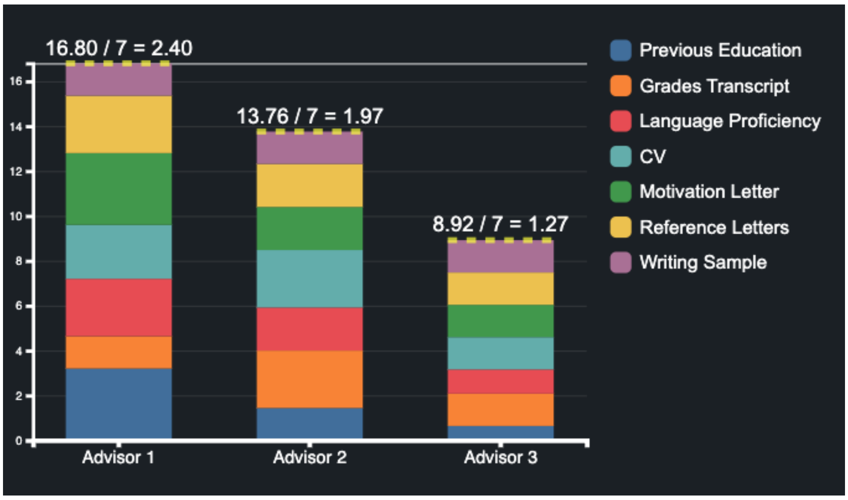
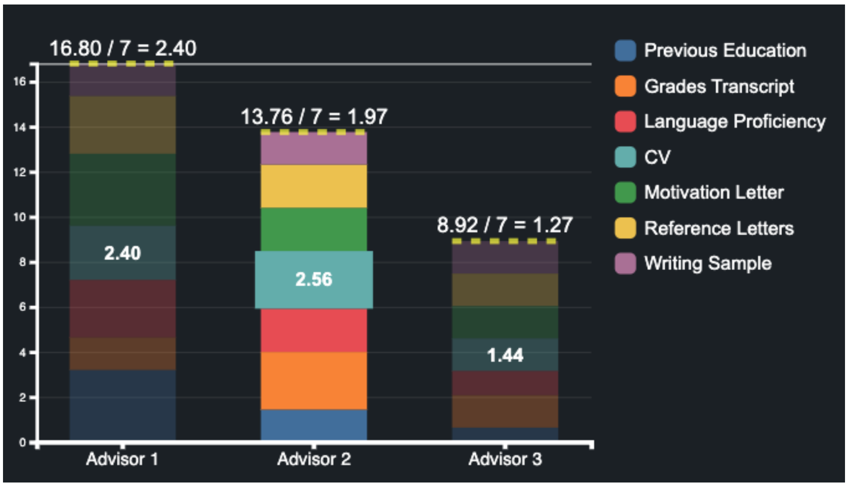


Figure 30: Hover effect of the Tabular Panel Bar Chart, showing additional comparison lines for clearer visual comparison between advisor ratings and application components.

Stack Bar Chart:
The Stacked Bar Chart (see Figure 31) offers a comparative view of the total scores across the seven application data points for each advisor, based on the selected filter—whether specific metrics or the composite score of each application component. By stacking the individual scores into a single bar for each advisor, the chart enables quick identification of which criteria had the greatest impact on an advisor's overall evaluation. It also facilitates comparison of the total evaluation between advisors. The layout highlights both the overall score and the relative contribution of each data point within that score. The use of distinct color blocks, consistent across all visualizations, allows users to easily track each application component's influence. Additionally, the total score is calculated and indicated at the top of each stack with the yellow line, making it easier to compare final composite values. The yellow line shows the sum of the values divided by the total number of application components (seven), ensuring transparency in the calculation. Both the sum and division are displayed directly in the chart to avoid hiding any information. If the chart is filtered by a single metric (such as weight or uncertainty), the average total value will only indicate that specific metric, while the application itself is evaluated based on all inputs and metrics. Thus, detailed calculations are displayed to ensure transparency in the evaluation process.



(a) Static view of the Stacked Bar Chart showing the total scores across seven application data points for each advisor. The visual marks used are rectangular bars, each representing an individual application component. Each stack's height encodes a value from 0 to 5 using the length channel, while the position channel along the x-axis represents individual advisors. The y-axis shows the cumulative scores of all application components. The same consistent color hue is used for each application component across the chart, with the total score displayed as the sum of scores divided by the total number of components, shown above the yellow line.



(b) Hover effect view of the Stacked Bar Chart, where additional details on the exact values are displayed. The hover interaction allows users to view precise scores for each component. Additionally, hovering over the legend highlights the corresponding values in the chart by adjusting the color and transparency of the highlighted blocks.

Figure 31: Stacked Bar Chart comparing the static view (a) with the hover effect (b), showing cumulative scores for each advisor. The x-axis represents the advisors, and the y-axis represents the total scores from 0 to 35, with individual application components stacked vertically. Composite values are highlighted at the top of each stack. The use of marks (rectangles) and channels (position, length, and color) aids in comparing both individual and overall scores across advisors.

Additionally, highlighting a value in the stacked bar chart also highlights the corresponding value in the Tabular Panel Bar Chart by displaying a yellow border, allowing for easy tracking of connected values (see Figure 32).

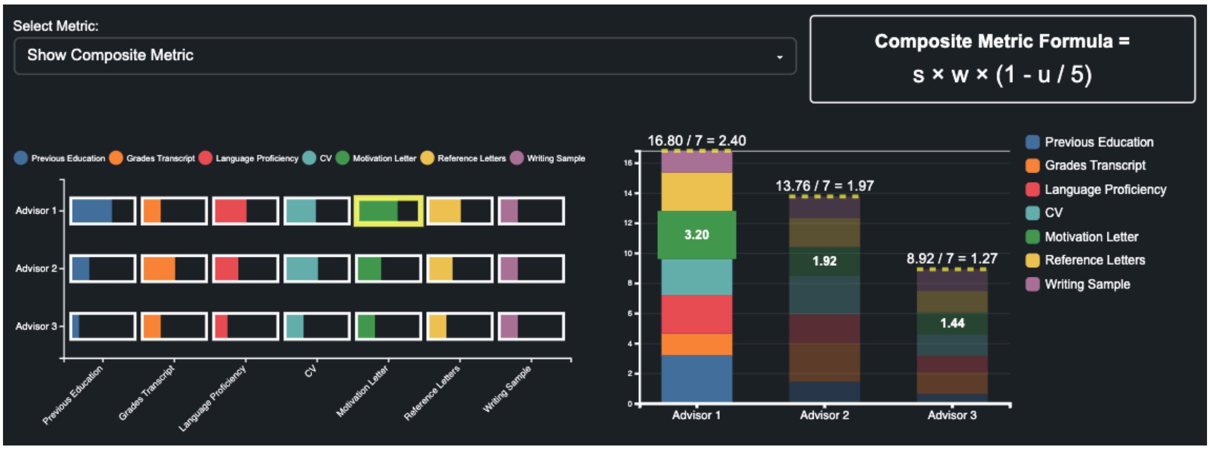
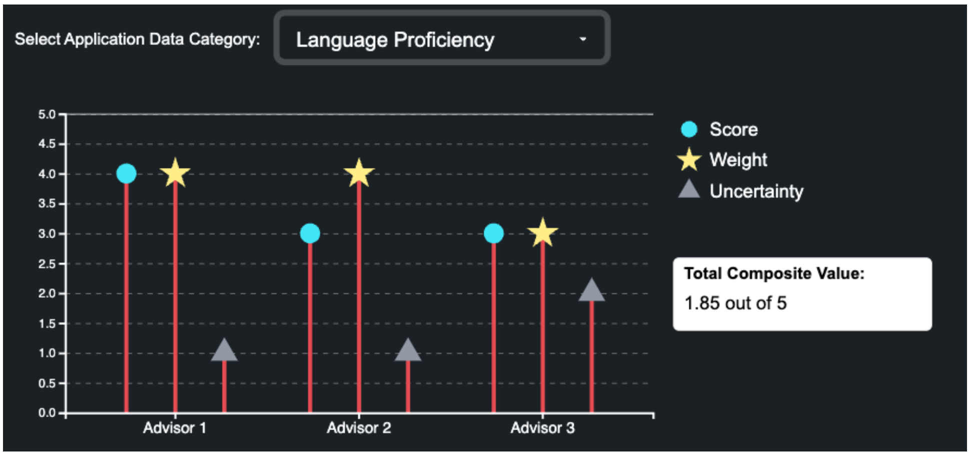


Figure 32: Interactive behavior showing that hovering over a value in the Stacked Bar Chart (right) highlights the corresponding value in the Tabular Panel Bar Chart (left) with a yellow border, allowing for visual tracking and comparison of related data points across both charts.

Lollipop Bar Chart:

The Lollipop Bar Chart (see Figure 33) is designed to provide a detailed breakdown of the Score, Weight, and Uncertainty assigned to each application data point by the advisors. Each metric is represented by the height of the line, with specific shapes and colors indicating Score (circle), Weight (star), and Uncertainty (triangle), effectively highlighting differences in these metrics across advisors. This chart emphasizes the magnitude of each metric, making it easier for evaluators to identify trends, disparities, or outliers. By presenting these metrics side by side, the visualization facilitates quick interpretation and comparison across advisors.



(a) Static view of the Lollipop Bar Chart showing the Score, Weight, and Uncertainty for each advisor. The visual marks used are shapes: circle for Score, star for Weight, and triangle for Uncertainty. The x-axis represents individual advisors, while the y-axis represents values from 0 to 5. The different shapes and consistent color coding (blue for Score, yellow for Weight, and gray for Uncertainty) help distinguish between the metrics.

(b) Hover effect view of the Lollipop Bar Chart, displaying additional details such as advisor comments for the selected application component. When hovering over a specific point, the related comment is shown on the right side of the chart. Hovering also highlights the corresponding score, weight, and uncertainty values for better comparison. Additionally, the total composite value is visible in the white box on the right.

Figure 33: Lollipop Bar Chart comparing the static view (a) with the interactive hover effect (b). The chart visualizes the Score, Weight, and Uncertainty for each advisor per application component, using position (on the y-axis) and shape to encode the metrics. The hover effect provides additional context, such as advisor comments and precise values for each metric.

Additionally, when hovering over specific points in all charts (excluding the summary chart), any comments provided by advisors for that particular application component are displayed, offering further qualitative insights into the evaluation. This visualization serves as a useful tool for detailed analysis, especially when evaluators need a closer look at how individual metrics are distributed across advisors.

Overall, the Visualization Section is designed to minimize cognitive load while maximizing the evaluators' ability to interpret raw, multi-dimensional data and derive combined quantitative values. The consistent use of color, interactive elements, and visual channels across the various charts ensures that users can transition seamlessly between different data representations. The decision not to highlight differences in advisors' input was intentional and aimed at encouraging the use of visualizations as conversation starters. By avoiding the emphasis on any particular input, the goal was to facilitate a more balanced discussion where all perspectives were equally considered. This approach allowed participants to engage with the data holistically, rather than focusing on discrepancies between individual assessments, ensuring that the visualizations supported a collaborative decision-making process without overemphasizing one advisor's input over another. Additionally, subjective and objective data were not separately visualized, despite initial consideration. The rationale behind this was the recognition that discussions around subjective assessments, such as personal motivations and recommendation letters, still require deeper deliberation and cannot be fully captured through visual tools. When rated metrics differed from one another, it was important for the group to engage in direct discussion to reconcile those differences. Since the data was simulated and participants were role-playing as advisors, it was not an ideal environment to perform a reliable evaluation of such nuanced subjective information. Thus, the study prioritized objective data for visualization while leaving space for discussion on subjective elements when conflicting assessments emerged.

The next section, Section 5, delves into the evaluation and findings from the study of EvaluationViz.

5 Evaluation

This section covers the evaluation of EvaluationViz, focusing on the qualitative role that visualizations played in the decision-making processes during the controlled experiment. In this study, participants were placed in a simulated environment where they assumed the role of admission officers, tasked with evaluating anonymized applicant profiles. The methods employed in the study are outlined in subsection 5.1, followed by a detailed analysis of the key findings in subsection 5.2, which compares participants' experiences in evaluating applicants both with and without visualizations.

5.1 Method

5.1.1 Participants & Recruitment

Participants for the EvaluationViz study were recruited through convenience sampling [17], leveraging both personal and professional networks. A total of twelve participants, divided into four groups of three, took part in the experiment. Each group was selected to ensure a mix of experience levels in application evaluation, enabling the study to explore how EvaluationViz performed with participants ranging from those with formal admissions experience to those without.

All participants were either extremely or slightly familiar with the field of information visualization, providing relevant context for their evaluations. Additionally, seven out of the twelve participants had real-life involvement with the admissions process, which enriched their ability to engage with the task. This combination allowed the study to gather insights from those familiar with admissions as well as from participants more accustomed to working with data visualization tools.

The sample consisted of five males, one non-binary individual, and six females, ranging in age from 30 to 49, with backgrounds in data analysis, UX design, human resources, and front-end engineering. Participants represented diverse nationalities, including three Germans, two Dutch, four individuals from Russia, Belarus, or Serbia, one Spanish, one Nepalese, and one Turkish participant. In this simulated university admissions environment, participants were tasked with evaluating anonymized applicant profiles, contributing to a controlled yet realistic decision-making setting.

5.1.2 Materials

The evaluation was conducted using a pre-loaded version of EvaluationViz on the author's laptop, accompanied by supplementary materials such as notebooks and pens for note-taking. Water was provided to participants during the session. The tool was presented in its final form, featuring two application datasets in JSON format for candidates applying to the HCI program, along with one program policy. The four primary visualizations—radar chart, tabular panel bar chart, stacked bar chart, and lollipop chart—were fully functional and ready for use during the evaluation.

5.1.3 Procedure

The evaluation was conducted in two distinct phases, with a four-day gap in between. Each phase consisted of two parts: an individual assessment followed by a group discussion. In Phase 1, participants evaluated applicants using only the provided forms, without the aid of visualizations, while Phase 2 introduced visualizations to support the decision-making process. Prior to each phase, participants were given a brief overview of the study's purpose and provided informed consent. However, detailed explanations were kept to a minimum to avoid influencing their natural thought processes and to encourage organic exploration of the tools.

In both phases, participants evaluated applicants based on seven key data points. The same evaluation rubric and criteria related to the HCI program were used in both phases to maintain consistency. Two comparable but slightly varied simulated applicant datasets were employed, allowing participants to assess one dataset in each phase. This controlled setup enabled direct comparisons between decision-making processes, both with and without the use of visual aids, and between individual and group assessments, while minimizing variability that could arise from using different application scenarios.

This approach was designed to align with the study's main research question and sub-questions, particularly those concerning how visualizations aid comprehensive assessment (RQ2), how they support collaborative decision-making (RQ3), and their role in reducing uncertainty and facilitating conversation during the final decision-making process (RQ4). Insights from expert interviews, conducted early in the study, highlighted that a typical admissions process begins with individual evaluations followed by group discussions. This informed the study's structure, allowing to examine how participants individually assessed applicants, and then came together to form group decisions, both with and without the use of visualizations. The focus was less on the final evaluation outcomes and more on understanding participants' thought processes during the evaluation, providing a deeper exploration of how visualizations influenced their decision-making behavior.

- Evaluation Phase 1:

The detailed questionnaire can be found in Appendix 4. It was designed to capture both individual and group decision-making processes, focusing on participants' reasoning, confidence levels, and reflections on the assessment tasks. The questions aimed to explore how participants made decisions without the aid of visualizations, as well as their experiences during group discussions.

Individual Experience:
In the first phase, participants were introduced to the evaluation form without any visualizations and asked to assess an anonymized applicant. Their task was to make an initial decision on whether to accept or reject the applicant, providing reasons for their choice, explaining how they formed their decision, and indicating their confidence levels via the post-assessment questionnaire.

Group Experience:
Additionally, during Phase 1, participants were asked to evaluate the same applicant as a group, again without the aid of visualizations. During the group discussion, they collectively decided on whether to accept or reject the applicant. Following the group session, participants completed the remainder of the questionnaire, which focused on their group experience, including the group's final decision, how the decision was reached, and how confident they were that all relevant factors had been considered.

- Evaluation Phase 2:
The questionnaire, detailed in Appendix 5, built upon the questions from Phase 1, incorporating participants' experiences with visualizations. It aimed to assess how the visual aids influenced both individual and group decision-making, specifically focusing on how participants used the visualizations, ranked their usefulness, and evaluated their trust in the visualized data.

Individual Experience:
In Phase 2, conducted four days later, participants were presented with a different anonymized applicant. This time, in addition to the evaluation form, participants were introduced to visualizations to aid in their decision-making. During the individual phase, participants could only see their own inputs visualized, without access to inputs from other advisors. They were then asked to decide whether to accept or reject the applicant and to describe how they used the visualizations to make their decision in the questionnaire.

Group Experience:
In the same Phase 2, participants re-evaluated the new applicant as a group, this time with the support of visualizations. Together, they discussed the visualized data and their insights to reach a final consensus on whether to accept or reject the applicant. Afterward, they completed the questionnaire, describing how the group used the visualizations, how the final decision was made, and ranking the visualizations based on their usefulness in facilitating the group's decision-making process.

At the end, participants were also presented with a demographic questionnaire, as detailed in Appendix 6 for research purposes. The complete process steps is also visualized in Figure 34.

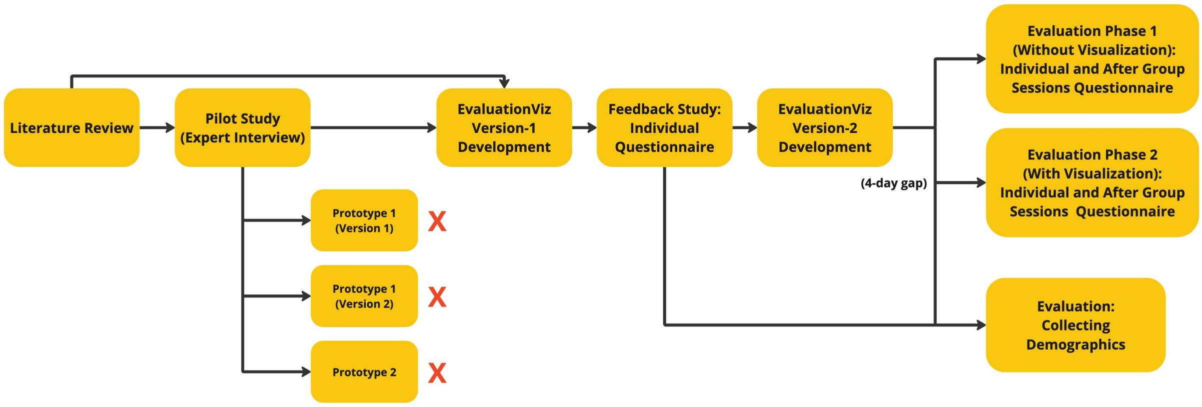


Figure 34: Workflow of the EvaluationViz Development and Evaluation Process. The process begins with a literature review and a pilot study involving expert interviews, which informed the initial design of the EvaluationViz tool. During development, several prototypes were created but were discontinued early as the thesis research question and scope became more refined. Two key versions of the EvaluationViz tool emerged: Version 1, developed before the feedback session, and Version 2, refined after incorporating insights from the feedback study. The final version was evaluated through two phases: Phase 1 (without visualizations) and Phase 2 (with visualizations), with a 4-day gap between the phases. Both phases involved individual and group sessions, followed by demographic data collection.

5.2 Findings

The analysis of the gathered data from the evaluation focused on observing how participants interacted with their simulated roles and how their thought processes and experiences differed with and without visualization aid. It also explored how the presence of visualizations influenced decision-making, particularly in terms of individual and group dynamics. Subsubsection 5.2.1 highlights the differences between the two phases, with participants reflecting on their experiences. Subsection 5.2.2 takes a deeper dive into analyzing the textual feedback provided by participants.

5.2.1 Contrasting Experiences: With and Without Visualization

The findings from both phases—without and with visualizations—revealed differences in how participants approached the decision-making process and how they explained the reasoning behind their decisions, as well as their confidence in considering all aspects, including application data and the metrics (weight, uncertainty) associated with each data point. In Phase 1, where participants evaluated applications and formed a decision without the support of visual aids, several participants reported challenges in synthesizing various elements of the application data, such as grades, alongside the metrics they were asked to assign and consider (scores, weights, and uncertainties). One participant noted, ”I reviewed each section of the applicant's data, but it is not straightforward to weigh them all together in the final decision. The grades seemed high enough, so I felt that should be a good indicator, but I wasn't sure how that was influenced by the uncertainty and weight I had assigned.” In general, participants tended to focus more on isolated aspects of the application data, particularly grades, without a clear method for incorporating metrics such as assigned uncertainty into their overall decision-making. Even when they acknowledged uncertainty, it was often based on their overall impression or uncertainty about the process, rather than a direct assessment of individual or multiple data points that contributed to the uncertainty.

In group discussions during Phase 1, while participants were able to reach consensus, their discussions often centered on the application data scores, such as grades or impressions on the motivation letter, without providing deeper reasoning or evidence to support their decisions. The group members rarely elaborated on why they assigned particular scores, weights, or uncertainties, and how they formed their decision including all points. One participant explained, "We started one by one explaining our opinions," but the discussions did not often extend to a more detailed analysis of the underlying reasons behind their evaluations while it was asked of them multiple times. This lack of evidence-based argumentation suggests that, despite having all the data in front of them while completing the questionnaire, participants tended to rely on surface-level metrics or the most immediately salient aspects of their evaluations, rather than synthesizing the information into a comprehensive analysis.

It is important to recognize that participants' ability to express their thought processes and articulate their reasoning may have been influenced by their proficiency in English, communication skills, and overall comfort in articulating complex ideas. This variation was evident in both phases of the study, as the way participants described their decision-making processes differed. In real-world admissions scenarios, where evaluators often come from diverse linguistic and cultural backgrounds, such factors could influence how they interpret and communicate insights from application data and assigned metrics. These challenges in articulation may have affected how clearly participants conveyed their reasoning, particularly when evaluating metrics within the application data. Positive experiences shared during the with-visualization phase revealed that compared to their experience in Phase 1 without visual aid, participants found it easier to back up their insights by first pointing at where they gave the most weight and where in data they had uncertainty in forming the final decision and initiate grounded discussions based on their evaluations and written comment per point.

In Phase 2, where visualizations such as the Radar Chart and Composite Score were introduced, participants generally reported greater clarity and structure in their evaluations. One participant noted, "The radar chart helped focus on the applicant's strengths and weaknesses, giving me more confidence in my decision." The visual tools provided participants with a more holistic view of the applicant's profile, incorporating not only grades but also the metrics such as scores, weightings, and uncertainties that were assigned by the advisors. All participants indicated that they began with the Composite Score and Radar Chart to get an initial overview before examining specific metrics within the application data in more detail.

Moreover, group discussions in Phase 2 appeared more organized, with visual tools serving as a reference point throughout the conversation. The visual aids offered a shared framework that seemed to help participants communicate their thoughts more effectively, contributing to more focused and cohesive discussions. As one participant reflected, "Because the visualizations present the data in a structured way, we could all quickly agree." The visual tools appeared to facilitate discussion by presenting multiple aspects of the application data evaluation inputs clearly, including metrics like scores and uncertainties. Overall, group discussions in Phase 2 tended to focus more on a comprehensive understanding of the applicant's profile, with participants referring to all relevant data points and metrics presented through the visualizations.

This observation directly relates to the research question about how visualizations can support collaboration and communication in group decision-making (RQ3). In real-world settings, where admissions officers may come from diverse backgrounds and have varying levels of experience and language proficiency, visualization can help bridge communication gaps. By offering a shared reference, visualizations may reduce misunderstandings and support more inclusive and transparent decision-making processes by also enabling the evaluator who is more familiar with the application background and has less uncertainty address the reasoning of their decision to accept the candidate. In this study, participants' descriptions of their thought processes in Phase 2 indicated a higher level of confidence and clarity, likely due to the structured nature of the visual aids.

In summary, the introduction of visualizations in Phase 2 appeared to enhance participants' confidence and their ability to engage with multiple application data points and metrics in a systematic way. The visual aids helped guide both individual assessments and group discussions, encouraging participants to consider a broader range of metrics beyond grades. However, it is also important to recognize that participants' ability to describe their thought processes may have been influenced by factors such as language proficiency and communication skills, which could impact decision-making in real-world admission settings. These findings underscore the value of visual tools in promoting clearer communication and more effective collaboration, particularly in diverse and international teams.

5.2.2 Insights and Interpretation of Collected Data

The qualitative findings from both phases of the study offer insights into how visualizations affected participants' decision-making processes, particularly in relation to RQ1. As mentioned in the previous subsection, in Phase 1, where participants evaluated applications without visual aids, they focused primarily on isolated metrics such as grades. Without the structure provided by visualizations, participants struggled to integrate additional metrics like assigned weights and uncertainties into their overall evaluation. As seen in the sentiment analysis [64] heatmap (Figure 35), participants reported higher levels of subjectivity in Phase 1, indicating a reliance on personal interpretation and intuition slightly more compared with visualization aid.

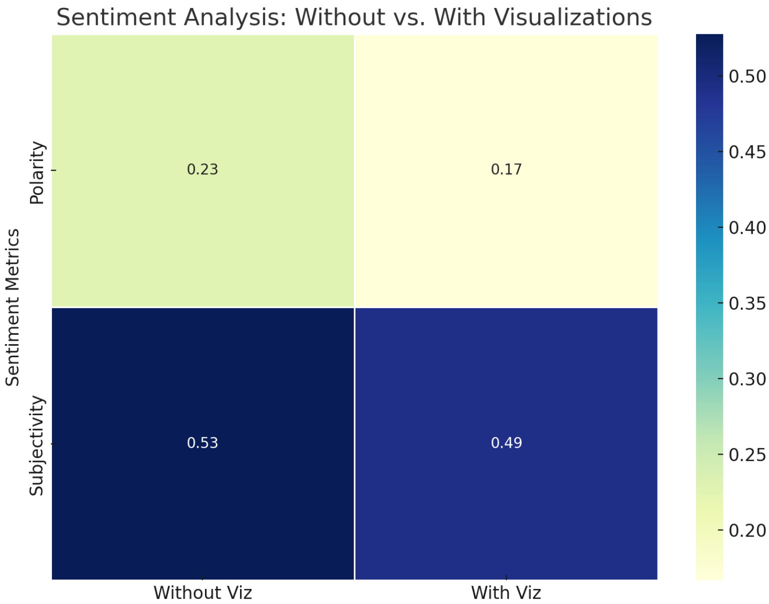


Figure 35: The heatmap compares two sentiment measures: polarity and subjectivity. Polarity reflects the overall positivity or negativity of responses, with values closer to 1 indicating positive sentiment and values closer to -1 indicating negative sentiment. Subjectivity measures how subjective or objective the responses are, with higher values indicating more personal, subjective responses, and lower values suggesting more objective, data-driven descriptions. The two columns represent sentiment scores from participants evaluating applications without visualization and with visualization tools. Without visual aids, the polarity score is 0.23, indicating mildly positive sentiment, whereas with visual aids, the score decreases slightly to 0.17, showing a marginal reduction in positivity. The subjectivity score without visual aids is 0.53, indicating more subjective interpretations of the decision-making process. This decreases to 0.49 with visual aids, suggesting that participants became slightly more objective in their evaluations when using visual tools, likely due to the structured nature of the data presented. This was calculated using Python's TextBlob library, which analyzes the textual responses to measure positivity and subjectivity levels based on the participants' descriptions in both phases [64].

In contrast, Phase 2, with the introduction of visual tools like the Radar Chart and Composite Score, shifted participants' focus to a more holistic evaluation, which relates to RQ2. Visualizations encouraged participants to systematically incorporate scores, weights, and uncertainties into their decision-making, resulting in more structured evaluations. The Radar Chart comparison (Figure 36) further illustrates this shift. Participants in Phase 2 demonstrated higher clarity and confidence, while levels of doubt decreased.

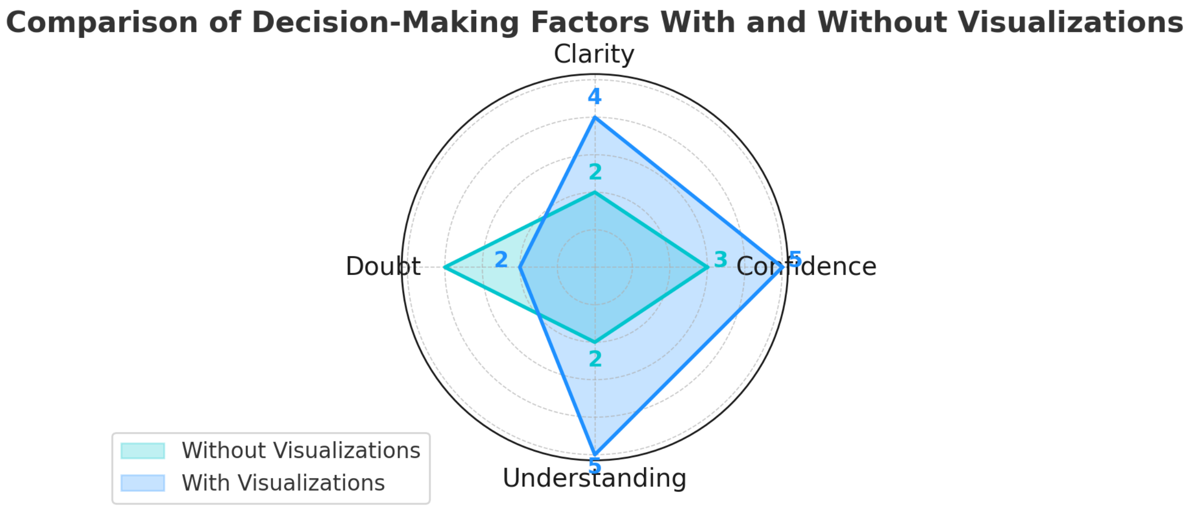


Figure 36: The chart was generated using Python's matplotlib library. Scores for key decision factors were derived from qualitative insights and thematic analysis from participant responses, comparing individual and group decisions across both phases [62]. The chart illustrates higher confidence and clarity, with reduced doubt, when visual aids were used in Phase 2.

In terms of the ranking provided by participants, the Radar Chart and Composite Score emerged as the most effective for both individual and group decision-making. Participants often pointed to the Radar Chart's ability to offer a clear, comprehensive view of applicants' strengths and weaknesses, integrating input from all evaluators alongside the overall composite score. This feature allowed for easy comparison against hypothetical thresholds (which participants themselves suggested, as the study did not include fixed thresholds) and facilitated the identification of key discussion points. As a result, it was seen as an accessible approach for both personal reflection and collaborative evaluation. Similarly, the Composite Score was appreciated for its streamlined summary of multiple metrics, helping to simplify complex data and guiding evaluators in prioritizing key factors (such as targeted differences between decisions) during the decision-making process.

Other visual tools, such as the Lollipop Chart and Tabular Panel Bar Chart, though ranked lower, were still recognized for their role in presenting structured data during group discussions. The Lollipop Chart, in particular, was valued for its ability to visually compare weighted metrics, though participants noted that it was more suited for detailed analysis. The Tabular Panel Bar Chart was perceived as less intuitive but still useful in overall comparison. While ranked lowest in terms of perceived usefulness, the stacked bar chart was acknowledged for providing a quick and effective overview of key information.

Moreover, In Phase 2, participants were asked how well the visualizations would support the group in recalling and understanding decisions at a later stage. The feedback was generally positive, with all participants suggesting that the visualizations, along with a comment section, could serve as useful reference points, particularly when decisions involved multiple evaluators or needed further justification.

These observations suggest that the introduction of visualization influenced participants' decision-making by fostering a more structured and transparent approach. The next chapter will dive deeper into the discussion 6 of these findings, exploring the impact of study limitations 6.1 and offering suggestions for future research 6.2, followed by the final chapter, Conclusion 7, which synthesizes and brings closure to the thesis.

6 Discussion

6.1 Limitation

While the EvaluationViz tool has demonstrated potential in enhancing decision-making in university admissions, the study conducted for this research has a few limitations that provide opportunities for future development. First, the use of simulated applicant profiles, rather than real-world educational data, was necessary to maintain ethical and privacy standards. However, real-world applications often include more complexities, such as varied grading systems, multiple transcripts, and a wider range of educational backgrounds. These nuances were not fully replicated in the simulated profiles, which may not have allowed participants to engage with the complete range of challenges faced in an actual admissions setting. This suggests that further research using real-world data could better test the tool's effectiveness and refine its functionalities to handle the complexities of authentic admissions decisions.

Additionally, the study took place in a controlled environment, which does not entirely replicate the high-pressure, high-stakes nature of real-world admissions processes. While this controlled setup allowed for clearer insights into the decision-making process, it may not have captured the full complexity of admissions dynamics, such as time constraints and the volume of applications typically encountered.

Participants, recruited through convenience sampling, were largely drawn from the researcher's professional network, which, though providing valuable insights, may have introduced some biases and limitations in terms of diversity of professional experience. As a result, the generalizability of the findings could be expanded with a larger and more varied sample in future studies, especially by incorporating a broader range of backgrounds and expertise in admissions.

Moreover, while this study focused primarily on quantitative metrics, less emphasis was placed on the qualitative components of the applications, such as motivation letters and reference letters. The visualizations were designed to represent data in a structured, quantifiable way, which may have led to the underrepresentation of these qualitative factors in participants' final decisions. Future iterations of the tool could better integrate qualitative data, ensuring that all aspects of an applicant's profile—both quantitative and qualitative—are comprehensively evaluated to support a more holistic assessment process.

Finally, there is a clear opportunity for future research to explore the tool's use in real-world settings. Conducting field studies with actual admissions professionals and using real educational history data would offer valuable insights into how EvaluationViz performs under real conditions. This would allow researchers to assess its practical utility in making high-stakes decisions and further refine the tool in collaboration with admissions experts. These considerations are proposed in the subsection 6.2, which outlines the next steps for advancing this research.

6.2 Future Work

There are several directions for future work that could address previous subsection mentioned limitations and enhance the capabilities of EvaluationViz. One promising avenue involves the integration of historical admissions data into the tool, allowing evaluators to compare current applicants with past applicants who had similar profiles. This could provide valuable insights into long-term trends and strategies, enabling evaluators to predict an applicant's future success based on patterns from previous admissions cycles. By connecting historical data with real-time evaluations, EvaluationViz could help admissions committees better understand which applicant characteristics are most predictive of success in their programs, thus supporting more informed and data-driven decisions.

Moreover, future iterations of the tool could benefit from automating certain aspects of the evaluation process. For example, automating the language proficiency and grade verification process using widely available educational data could reduce the workload for evaluators and increase the accuracy of the initial screening. Natural language processing (NLP) could be employed to assess language proficiency by analyzing writing samples while grading data could be automatically normalized and compared across different educational systems. This automation would streamline the decision-making process, allowing human evaluators to focus on more subjective aspects of the application, such as motivation letters or reference letters, which cannot yet be fully automated as it requires human judgment to fully evaluate students' personalities and fit. These processes could also be enhanced by real-time visualization, allowing evaluators to instantly see how an applicant's grades and language proficiency compare to a predefined benchmark or similar candidates.

Another exciting possibility is the use of machine learning algorithms to further assist in the admissions process. By training models on historical admissions and performance data, the tool could offer evaluators suggestions based on patterns in the data. For instance, it could highlight candidates who have profiles similar to previously successful applicants or flag those who are outliers in certain metrics. These insights could be integrated into the visualizations to provide a more nuanced understanding of each applicant's potential. This approach could also be extended to incorporate feedback loops, where evaluators' decisions are tracked over time to continuously refine and improve the model's recommendations.

The evolution of the tool could also be aligned with developments in collaborative decision-making frameworks. While the current focus is on enhancing individual and small-group decision-making, future versions could incorporate features that support larger admissions committees. For example, adding features such as comment threads, voting mechanisms, and consensus-building tools could help larger groups of evaluators work together more efficiently. Furthermore, the tool could evolve to track the decision-making process over time, providing insights into how and why certain decisions were made, and offering a record for transparency and accountability.

Finally, usability studies should continue to refine the design and functionality of the tool. Although the feedback from the current study was largely positive, future work should focus on increasing the flexibility and adaptability of the tool, ensuring that it can cater to the specific needs of different institutions. Customization options, such as allowing institutions to define their own metrics and criteria for evaluation, could be implemented to make the tool more versatile and applicable to a wider range of admissions processes.

7 Conclusion

In conclusion, EvaluationViz offers a promising approach to improving transparency, consistency, and collaboration in university admissions through structured decision-making frameworks. The tool's ability to consolidate key metrics—such as score, weight, and uncertainty—across multiple data points allowed evaluators to assess applicants in a more holistic and confident manner. Visual aids, such as the Radar Chart, Lollipop

Chart, and Tabular Panel Bar Chart, helped clarify and prioritize data, making it easier to compare different applicant attributes and reach well-supported decisions. The author suggests further research on integrating visual encodings, such as Sankey diagrams and incorporating more dynamic subjective dimensions, such as leadership skills and communication skills. Additionally, the inclusion of applicant history, showing the tracking of similar students, could help make more informed decisions on the evaluation of new applicants while ensuring proper transparency in the decision-making process.

The development of EvaluationViz was guided by expert interviews conducted early in the research process, which provided valuable insights into how admissions committees currently make decisions and the challenges they face when evaluating multiple criteria. This feedback, addressing RQ1, revealed the limitations of current tools and the subjectivity often present in decision-making, leading to the design of EvaluationViz to better support admissions officers by offering a more comprehensive and structured view of applicants.

The findings from the evaluation study suggest that the proposed solution has a positive influence on both individual and collaborative decision-making, directly addressing RQ2, RQ3, RQ4, and the main research question. Participants in the evaluation reported increased confidence in their decisions, especially in Phase 2, where visualizations helped them better synthesize all data points. The Radar Chart was particularly noted for providing a clear overview of applicant strengths and weaknesses, facilitating more data-driven decisions. Additionally, the focus group sessions demonstrated that visual tools helped align evaluators' judgments more effectively, leading to quicker and more transparent consensus within the groups. This aligns with RQ3, as the visualizations offered a shared framework that encouraged clearer communication and recall of key data points during discussions.

However, the study also highlights areas for future improvement. The use of simulated data and a controlled environment, while necessary for initial testing, limits the generalizability of the findings. Real-world admissions involve more complexities, such as varied educational backgrounds and qualitative data, which were not fully captured in this study. This presents an opportunity for future research to explore how visualizations can integrate both quantitative and qualitative factors, addressing RQ4. Participants indicated that while the visual tools supported structured data analysis, qualitative elements, such as motivation letters and personal statements, were less emphasized. Future iterations of the tool should aim to better incorporate these qualitative aspects to ensure a more holistic assessment of applicants.

Looking ahead, EvaluationViz holds potential. By integrating historical data, enhancing automation, and improving collaborative features, the tool could become an invaluable asset for admissions committees. Addressing the current limitations and expanding its capabilities will allow EvaluationViz to support more informed, fair, and efficient admissions decisions. Additionally, participants noted that having access to historical data could provide valuable insights, allowing comparisons between current applicants and past successful candidates, and potentially offering predictive insights into future performance.

References

[1] Amjad Alowaigl, Khalil Al-Shqeerat, and Mohammed Hadwan. 2020. A multi-criteria assessment of decision support systems in educational environments. *Indonesian Journal of Electrical Engineering and Computer Science* 22, 2 (2020), 377–388. https://doi.org/10.11591/ijeecs.v22.i2.pp985-996

[2] Common App. 2024. Common App Deadline Updates 2024. https://www.commonapp.org/files/Common-App-Deadline-Updates-2024.03.01.pdf. Accessed: 2024-05-10.

[3] Giuseppe Carenini and John Loyd. 2004. Valuecharts: analyzing linear models expressing preferences and evaluations. In *Proceedings of the working conference on Advanced visual interfaces*. 150–157. https://doi.org/10.1145/989863.989885

[4] Lena Cibulski, Evanthia Dimara, Setia Hermawati, and Jörn Kohlhammer. 2022. Supporting domain characterization in visualization design studies with the critical decision method. In *2022 IEEE 4th Workshop on Visualization Guidelines in Research, Design, and Education (VisGuides)*. IEEE, 8–15. https://doi.org/10.1109/VisGuides57787.2022.00007

[5] William S Cleveland. 1993. A model for studying display methods of statistical graphics. *Journal of Computational and Graphical Statistics* 2, 4 (1993), 323–343. https://doi.org/10.1080/10618600.1993.10474616

[6] Cristina Conati, Giuseppe Carenini, Enamul Hoque, Ben Steichen, and Dereck Toker. 2014. Evaluating the impact of user characteristics and different layouts on an interactive visualization for decision making. In *Computer Graphics Forum*, Vol. 33. Wiley Online Library, 371–380. https://doi.org/10.1111/cgf.12393

[7] Stanford Daily. 2023. Acceptance rate drops to 3.68%, majority women and non-white students. https://stanforddaily.com/2023/02/28/acceptance-rate-drops-to-3-68-majority-women-and-non-white-students/ Accessed: August 31, 2024.

[8] Aritra Dasgupta, Susannah Burrows, Kyungsik Han, and Philip J Rasch. 2017. Empirical analysis of the subjective impressions and objective measures of domain scientists' visual analytic judgments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 1193–1204. https://doi.org/10.1145/3025453.302588

[9] Marco Dean. 2020. Multi-criteria analysis. In *Standard Transport Appraisal Methods*, Niek Mouter (Ed.). Advances in Transport Policy and Planning, Vol. 6. Academic Press, 165–224. https://doi.org/10.1016/bs.atpp.2020.07.001

[10] Evanthia Dimara, Anastasia Bezerianos, and Pierre Dragicevic. 2017. Conceptual and methodological issues in evaluating multidimensional visualizations for decision support. *IEEE transactions on visualization and computer graphics* 24, 1 (2017), 749–759. https://doi.org/10.1109/TVCG.2017.2745138

[11] Evanthia Dimara, Steven Franconeri, Catherine Plaisant, Anastasia Bezerianos, and Pierre Dragicevic. 2018. A task-based taxonomy of cognitive biases for information visualization. *IEEE transactions on visualization and computer graphics* 26, 2 (2018), 1413–1432. https://doi.org/10.1109/TVCG.2018.2872577

[12] Times Higher Education. 2024. Utrecht University in Times Higher Education Rankings 2024. https://www.timeshighereducation.com/world-university-rankings/utrecht-university Accessed: August 31, 2024.

[13] AA Eliman. 1991. A decision support system for Univeristy admission policies. *European Journal of Operational Research* 50, 2 (1991), 140–156. https://doi.org/10.1016/0377-2217(91)90237-P

[14] Martin J. Eppler. 2021. The effects of visualization on judgment and decision-making. *Management Review Quarterly* 71 (2021), 399–428. https://doi.org/10.1007/s11301-021-00235-8

[15] Mastermind Europe. 2017. Admissions to English-Taught Programs (ETPs) at Master's Level in Europe – Procedures, Regulations, Success Rates, and Challenges for Diverse Applicants. https://mastermindeurope.eu/wp-content/uploads/2017/01/Report-2-Admissions-to-ETPs.pdf Accessed: August 31, 2024.

[16] Mastermind Europe. 2018. Coherent Admission Framework. https://mastermindeurope.eu/wp-content/uploads/2018/05/GT1-Coherent-Admission-Framework_May_2018.pdf Accessed: 2024-08-26.

[17] Farahman Farrokhi and Asgar Mahmoudi-Hamidabad. 2012. Rethinking convenience sampling: Defining quality criteria. *Theory & Practice in Language Studies (TPLS)* 2, 4 (2012). https://doi.org/10.4304/tpls.2.4.784-792

[18] Amy Franklin, Swaroop Gantela, Salsawit Shifarraw, Todd R Johnson, David J Robinson, Brent R King, Amit M Mehta, Charles L Maddow, Nathan R Hoot, Vickie Nguyen, et al. 2017. Dashboard visualizations: Supporting real-time throughput decision-making. *Journal of biomedical informatics* 71 (2017), 211–221. https://doi.org/10.1016/j.jbi.2017.05.024

[19] Brigid Freeman. 2020. Admissions Processes to Higher Education, International Insights. In *The International Encyclopedia of Higher Education Systems and Institutions*. Springer, 94–100. https://doi.org/10.1007/978-94-017-8905-9_341

[20] Catherine D Gamper and Catrinel Turcanu. 2007. On the governmental use of multi-criteria analysis. *Ecological economics* 62, 2 (2007), 298–307. https://doi.org/10.1016/j.ecolecon.2007.01.010

[21] Judy Wawira Gichoya, Kaesha Thomas, Leo Anthony Celi, Nabile Safdar, Imon Banerjee, John D Banja, Laleh Seyyed-Kalantari, Hari Trivedi, and Saptarshi Purkayastha. 2023. AI pitfalls and what not to do: mitigating bias in AI. *The British Journal of Radiology* 96, 1150 (2023), 20230023. https://doi.org/10.1259/bjr.20230023

[22] Leo A Goodman. 1961. Snowball sampling. *The annals of mathematical statistics* (1961), 148–170.

[23] Samuel Gratzl, Alexander Lex, Nils Gehlenborg, Hanspeter Pfister, and Marc Streit. 2013. Lineup: Visual analysis of multi-attribute rankings. *IEEE transactions on visualization and computer graphics* 19, 12 (2013), 2277–2286. https://doi.org/10.1109/TVCG.2013.173

[24] Patricia Gurin, Eric Dey, Sylvia Hurtado, and Gerald Gurin. 2002. Diversity and higher education: Theory and impact on educational outcomes. *Harvard educational review* 72, 3 (2002), 330–367. https://doi.org/10.17763/haer.72.3.01151786u134n051

[25] Gregory L Heileman, Terry H Babbitt, and Chaouki T Abdallah. 2015. Visualizing student flows: Busting myths about student movement and success. *Change: The Magazine of Higher Learning* 47, 3 (2015), 30–39. https://doi.org/10.1080/00091383.2015.1031620

[26] Andrew Herman and Mike Oliver. 2024. Alma SIS Pricing, Alternatives & More 2024. https://www.capterra.com/p/133283/Alma/

[27] Emily Hindalong. 2018. *Visual techniques for exploring alternatives and preferences in Group Preferential Choice*. Ph.D. Dissertation. University of British Columbia. https://doi.org/10.14288/1.0363067

[28] Emily Hindalong, Jordon Johnson, Giuseppe Carenini, and Tamara Munzner. 2020. Towards Rigorously Designed Preference Visualizations for Group Decision Making. In *2020 IEEE Pacific Visualization Symposium (PacificVis)*. 181–190. https://doi.org/10.1109/PacificVis48177.2020.5111

[29] Jimmy Johansson and Camilla Forsell. 2015. Evaluation of parallel coordinates: Overview, categorization and guidelines for future research. *IEEE transactions on visualization and computer graphics* 22, 1 (2015), 579–588. https://doi.org/10.1109/TVCG.2015.2466992

[30] Slava Kalyuga. 2011. Cognitive load theory: How many types of load does it really need? *Educational psychology review* 23 (2011), 1–19. https://doi.org/10.1007/s10648-010-9150-7

[31] Sami Koivunen, Thomas Olsson, Ekaterina Olshannikova, and Aki Lindberg. 2019. Understanding decision-making in recruitment: Opportunities and challenges for information technology. *Proceedings of the ACM on human-computer interaction* 3, GROUP (2019), 1–22. https://doi.org/10.1145/3361123

[32] Kees Kouwenaar. 2015. Changing paradigms: towards competency-assessment in admission to master's programs in Europe: a review. *Tuning J. High. Educ.* (dec 2015). https://doi.org/10.18543/tjhe-3(1)-2015pp99-135

[33] Kostiantyn Kucher, Carita Paradis, and Andreas Kerren. 2018. The state of the art in sentiment visualization. In *Computer graphics forum*, Vol. 37. Wiley Online Library, 71–96. https://doi.org/10.1111/cgf.13217

[34] Anastasia Kurysheva, Harold VM van Rijen, and Gönül Dilaver. 2019. How do admission committees select? Do applicants know how they select? Selection criteria and transparency at a Dutch University. *Tertiary education and management* 25, 4 (2019), 367–388.

[35] Heidi Lam, Enrico Bertini, Petra Isenberg, Catherine Plaisant, and Sheelagh Carpendale. 2011. Seven guiding scenarios for information visualization evaluation. *IEEE Transactions on Visualization and Computer Graphics* 18, 9 (2011), 1520–1536. https://doi.org/10.1109/TVCG.2011.279

[36] Cecilia Larsson. 2022. Cognitive Load in Advanced Systems: Registration of Swedish grades in the University Admissions process.

[37] Kelin Li, Jipeng Li, Yujing Sun, Chenhui Li, and Changbo Wang. 2023. Color assignment optimization for categorical data visualization with adjacent blocks. *Journal of Visualization* 26, 4 (2023), 917–936. https://doi.org/10.1007/s12650-022-00905-z

[38] Qianyu Liu, Haoran Jiang, Zihao Pan, Qiushi Han, Zhenhui Peng, and Quan Li. 2024. BiasEye: A Bias-Aware Real-time Interactive Material Screening System for Impartial Candidate Assessment. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*. 325–343. https://doi.org/10.1145/3640543.3645166

[39] Steffen Lohmann, Florian Heimerl, Fabian Bopp, Michael Burch, and Thomas Ertl. 2015. Concentri cloud: Word cloud visualization for multiple text documents. In *2015 19th International Conference on Information Visualisation*. IEEE, 114–120. https://doi.org/10.1109/iV.2015.30

[40] Harvard Magazine. 2024. Admissions Statistics for the Class of 2024. https://www.harvardmagazine.com/2024/05/hj-admissions-class-of-2024 Accessed: August 31, 2024.

[41] 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*. 17–20. https://doi.org/10.1145/3022198.3023269

[42] R. Michel, S. Pollard, M. E. Oliveri, and C. Wendler. 2020. Higher Education Admissions Practices: An International Perspective. In *Higher Education Admissions Practices: An International Perspective*. Cambridge University Press, Chapter 1, 5–17. https://doi.org/10.1017/9781108559607.001

[43] Miro. 2024. Miro: Online Whiteboard for Visual Collaboration. https://miro.com/ Accessed: September 1, 2024.

[44] M. D. Mumford. 2013. Common Visualizations: Their Cognitive Utility. In *Handbook of Systems Engineering and Risk Management in Control Systems, Communication, Space Technology, Missile Technology and Process Control Systems*. Springer, 301–318. https://doi.org/10.1007/978-1-4614-7485-2_27

[45] Tamara Munzner. 2009. A Nested Model for Visualization Design and Validation. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 921–928. https://doi.org/10.1109/TVCG.2009.111

[46] Tamara Munzner. 2014. The Nested Model for Visualization Design and Validation. In *Visualization Analysis and Design*. A K Peters/CRC Press, New York, 13–34. https://doi.org/10.1201/b17511

[47] Arpit Narechania, Adam Coscia, Emily Wall, and Alex Endert. 2021. Lumos: Increasing awareness of analytic behavior during visual data analysis. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 1009–1018. https://doi.org/10.1109/TVCG.2021.3114827

[48] Nuffic. 2022-2023. Incoming Degree Mobility in Dutch Higher Education 2022-2023. https://www.nuffic.nl/sites/default/files/2023-06/incoming-degree-mobility-in-dutch-higher-education-2022-2023.pdf. Accessed: 2024-05-10.

[49] Nuffic. 2023. Factsheet International Students 2023. https://www.nuffic.nl/sites/default/files/2023-05/factsheet-international-students-2023.pdf. Accessed: 2024-05-10.

[50] Emre Oral, Ria Chawla, Michel Wijkstra, Narges Mahyar, and Evanthia Dimara. 2023. From information to choice: A critical inquiry into visualization tools for decision making. *IEEE Transactions on Visualization and Computer Graphics* (2023). https://doi.org/10.1109/TVCG.2023.3326593

[51] Oxford English Dictionary. [n.d.]. Admission. https://www.oed.com/view/Entry/2567. Accessed: 2024-09-06.

[52] Lindsey M. Padilla, Karen K. Schloss, Sarah H. Creem-Regehr, and Jeanine K. Stefanucci. 2022. Uncertainty Visualization. In *Handbook of Human-Computer Interaction*. CRC Press, Chapter 22, 519–548. https://doi.org/10.1201/9781003021339-22

[53] Stephan Pajer, Marc Streit, Thomas Torsney-Weir, Florian Spechtenhauser, Torsten Möller, and Harald Piringer. 2016. Weightlifter: Visual weight space exploration for multi-criteria decision making. *IEEE transactions on visualization and computer graphics* 23, 1 (2016), 611–620. https://doi.org/10.1109/TVCG.2016.2598589

[54] Lawrence A Palinkas, Sarah M Horwitz, Carla A Green, Jennifer P Wisdom, Naihua Duan, and Kimberly Hoagwood. 2015. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and policy in mental health and mental health services research* 42 (2015), 533–544. https://doi.org/10.1007/s10488-013-0528-y

[55] UK Parliament. 2023. International and EU students in higher education in the UK FAQs. https://commonslibrary.parliament.uk/research-briefings/cbp-7976/ Accessed: 2024-06-10.

[56] Arif Perdana, Alastair Robb, and Fiona Rohde. 2019. Interactive data and information visualization: unpacking its characteristics and influencing aspects on decision-making. *Pacific Asia Journal of the Association for Information Systems* 11, 4 (2019), 4. https://doi.org/10.17705/1pais.11404

[57] Luxsanan Ploywattanawong. 2016. Decision support systems model for admission to Higher Education. In *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*. IEEE, 1359–1363. https://doi.org/10.1109/CompComm.2016.7924926

[58] World Population Review. 2024. Grading Systems by Country 2024. https://worldpopulationreview.com/country-rankings/grading-systems-by-country. Accessed: 2024-05-12.

[59] Gretchen W Rigol. 2003. Admissions Decision-Making Models: How US Institutions of Higher Education Select Undergraduate Students. *College Entrance Examination Board* (2003).

[60] Bahador Saket, Alex Endert, and Çağatay Demiralp. 2018. Task-based effectiveness of basic visualizations. *IEEE transactions on visualization and computer graphics* 25, 7 (2018), 2505–2512. https://doi.org/10.1109/TVCG.2018.2829750

[61] Alper Sarikaya and Michael Gleicher. 2017. Scatterplots: Tasks, data, and designs. *IEEE transactions on visualization and computer graphics* 24, 1 (2017), 402–412. https://doi.org/10.1109/TVCG.2017.2744184

[62] Ali Hassan Sial, Syed Yahya Shah Rashdi, and Abdul Hafeez Khan. 2021. Comparative analysis of data visualization libraries Matplotlib and Seaborn in Python. *International Journal* 10, 1 (2021), 277–281. https://doi.org/10.30534/ijatcse/2021/391012021

[63] Herbert A. Simon. 1960. The new science of management decision. *Harper & Brothers* (1960). https://doi.org/10.1037/13978-000

[64] Amir Sohel, Muhammad Rabiul Hossain, Zubiya Binta Mostofa, Md Umaid Hasan, Utpal Chandra Das, and Sara Kifayath Parvin. 2023. Sentiment Analysis Based on Online Course Feedback Using Textblob and Machine Learning Techniques. In *2023 26th International Conference on Computer and Information Technology (ICCIT)*. IEEE, 1–6. https://doi.org/10.1109/ICCIT60459.2023.10441476

[65] Gaurav Somani. 2024. MasterSoft Online Admission System Pricing, Alternatives & More 2024. https://www.capterra.com/p/251942/Online-Admission-System/

[66] Anselm Strauss and Juliet Corbin. 1994. Grounded theory methodology: An overview. In *Handbook of qualitative research*, Norman K. Denzin and Yvonna S. Lincoln (Eds.). Sage Publications, Thousand Oaks, CA, 273–285.

[67] Studielink. 2024. Studielink: The official website for higher education applications in the Netherlands. https://www.studielink.nl/ Accessed: September 7, 2024.

[68] Justin Talbot, Vidya Setlur, and Anushka Anand. 2014. Four experiments on the perception of bar charts. *IEEE transactions on visualization and computer graphics* 20, 12 (2014), 2152–2160. https://doi.org/10.1109/TVCG.2014.2346320

[69] Kira Talent. [n.d.]. Kira Talent: The world's only holistic admissions solution for higher education. https://www.kiratalent.com/

[70] Kira Talent. 2018. Breaking down bias in admissions: The how-to guide to preventing admissions bias at your school. http://start.kiratalent.com/breaking-down-admissions-bias/ Accessed: 2024-09-05.

[71] 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), 1–22. https://doi.org/10.1145/3274438

[72] Ahmad Tariq and Khan Rafi. 2012. Intelligent decision support systems-A framework. In *Information and Knowledge Management*, Vol. 2. 12–20.

[73] The Straits Times. 2024. South Korea's International Student Population Jumps 10-Fold in 2 Decades. *The Straits Times* (mar 2024). https://www.straitstimes.com/asia/east-asia/south-korea-s-international-student-population-jumps-10-fold-in-2-decades#:~:text=Though%20the%20trend%20was%20stagnant,to%20181%2C842%20as%20at%202023.

[74] Alfred Ultsch and Jörn Lötsch. 2022. Euclidean distance-optimized data transformation for cluster analysis in biomedical data (EDOtrans). *BMC bioinformatics* 23, 1 (2022), 233. https://doi.org/10.1186/s12859-022-04769-w

[75] Radboud University. 2019. Conversion of Grades and Credits. https://www.ru.nl/sites/default/files/2023-05/mso-io_2016_028_conversion_grades_and_credits_radboud_2019_update.pdf Accessed: 2023-12-09.

[76] Utrecht University. [n.d.]. Graduate School of Life Sciences. https://www.uu.nl/en/education/graduate-school-of-life-sciences Accessed: September 4, 2024.

[77] Utrecht University. [n.d.]. Graduate School of Natural Sciences. https://www.uu.nl/en/organisation/graduate-school-of-natural-sciences Accessed: September 4, 2024.

[78] Yunhai Wang, Fubo Han, Lifeng Zhu, Oliver Deussen, and Baoquan Chen. 2017. Line graph or scatter plot? automatic selection of methods for visualizing trends in time series. *IEEE transactions on visualization and computer graphics* 24, 2 (2017), 1141–1154. https://doi.org/10.1109/TVCG.2017.2653106

[79] Susan C Weller, Ben Vickers, H Russell Bernard, Alyssa M Blackburn, Stephen Borgatti, Clarence C Gravlee, and Jeffrey C Johnson. 2018. Open-ended interview questions and saturation. *PloS one* 13, 6 (2018), e0198606. https://doi.org/10.1371/journal.pone.0198606

[80] Wanvipa Wongvilaisakul, Paniti Netinant, and Meennapa Rukhiran. 2023. Dynamic multi-criteria decision making of graduate admission recommender system: AHP and fuzzy AHP approaches. *Sustainability* 15, 12 (2023), 9758. https://doi.org/10.3390/su15129758

[81] Xiaojiao Xie, Fanghao Song, Yan Liu, Shurui Wang, and Dong Yu. 2021. Study on the effects of display color mode and luminance contrast on visual fatigue. *IEEE access* 9 (2021), 35915–35923. https://doi.org/10.1109/ACCESS.2021.3061770

1 Appendix A: Pilot Study Expert Interview Consent & Procurement Survey

Utrecht University

Welcome to my HCI Master Thesis research study!
Expert Interview in Utrecht University

You are invited to participate in an interview as part of a Master's Thesis research project in the field of **Human-Computer Interaction**, titled "**Visualizing Multi-Criteria Evaluations of Application Data in University Admissions: Supporting Holistic and Collaborative Decision-Making**." Based on your preference, the interview will be conducted online or in person in one of the available meeting rooms at the Buys Ballot Building (BBG) or Minnaert Building. This interview is essential for gaining insights into **real-world admission processes** and understanding how they work in various contexts. Your experience in the admissions process at **Utrecht University** can help provide valuable insights into common challenges and practices. The goal is to identify these challenges and gather input for designing a data visualization tool that improves transparency and reduces cognitive load in admissions decision-making. The interview is expected to last approximately **one hour**, and as a token of appreciation for your participation, refreshments and snacks will be provided.

Read the information below to understand the interview process, data collection methods, your rights, and the study's approval details for this research project:

- Interview Conductor
This interview is conducted by **Samin Asnaashari** as part of my Master's thesis in Human-Computer Interaction under the supervision of **Evanthia Dimara** (e.dimara@uu.nl) and **Christof van Nimwegen** (c.vanNimwegen@uu.nl)

- Study Procedure
The interview aims to gather insights into the university admissions process, with a focus on identifying challenges and uncertainties. It will contribute to the development of a visualization solution designed to support admissions evaluation and decision-making. The session will last approximately one hour.

- Data Handling
With your consent, the interview will be **audio-recorded and stored securely on Yoda** (a research data management service maintained by Utrecht University), accessible only by the researcher and thesis supervisor. The recording will be transcribed anonymously and deleted within 3 months. All provided information, including demographic details and opinions, will be anonymized or pseudonymized, ensuring your privacy. Your identifiable personal data will be deleted after analysis and replaced with an ID for data linkage.

- Your Rights
Participation is voluntary and based on your consent. You may withdraw at any time

without any obligation to justify your decision. Upon withdrawal, all personal data collected until that point will be erased, excluding anonymized data already processed.

- Study Approval
The Research Institute of Information and Computing Sciences has approved this study following an Ethics and Privacy Quick Scan. For complaints about the study conduct, contact ics-ethics@uu.nl. For personal data concerns or exercising GDPR rights, reach out to privacy-beta@uu.nl. Visit www.uu.nl/en/organisation/privacy for more on privacy policies and your data rights.

By signing this form, I acknowledge and consent to the following terms for my participation in the research project titled "**Visualizing Multi-Criteria Evaluations of Application Data in University Admissions: Supporting Holistic and Collaborative Decision-Making**":

- **Age Confirmation:**
  - I confirm that I am 18 years of age or older.
- **Project Understanding:**
  - I confirm that the research project has been clearly explained to me. I have had the opportunity to ask questions and have received satisfactory answers. I understand that I can ask further questions at any time and that I have enough time to consider my participation.
- **Consent to Use of Material:**
  - I consent to the material I contribute being used to generate insights for the research project.
- **Collection and Confidentiality of Personal Data:**
  - I consent to the use of audio recordings in this study as detailed above. I understand I have the right to request the deletion of recordings at any time.
  - I understand that if I give permission, the audio recordings will be accessed solely by **Samin Asnaashari.** The recordings will be transcribed, stored securely on the researcher's personal University drive, and will be deleted 3 months post-study. Transcriptions will be anonymized. In accordance with GDPR, I can access my recordings and request their deletion at any time during this period.
  - I understand that, in addition to recordings, other personal data collected from me will be held confidentially, with only **Samin Asnaashari** having access. This information will be stored securely and anonymized 3 months post-study. In accordance with GDPR, I can access and request the deletion of my data at any time during this period.
- **Voluntary Participation and Withdrawal Rights:**
  - I understand my participation is voluntary, and I am free to withdraw at any time without needing to provide a reason. Upon withdrawal, any personal data collected from me will be erased.
- **Consent for Anonymized Data Use:**
  - I consent to allow the fully anonymized data to be used in future publications and scholarly dissemination of the research findings.
  - I understand the data acquired will be securely stored, and appropriately anonymized data may be made available to others for research purposes.

The University may publish anonymized data in suitable data repositories for verification and to facilitate access by researchers and other users.

- **Data Deletion Request:**
  - I understand I have the right to request the deletion of any personal data collected from me.

◯ I consent, begin the study

◯ I do not consent, I do not wish to participate

Q: Please provide your **Full Name**:

Q: Please provide your **Email Address** for contact:

Q: What is the Age range to which you belong?
◯ 25-34 years old
◯ 35-44 years old
◯ 45-54 years old
◯ 55-64 years old
◯ 65+ years old

Q: Could you briefly describe your role in the admission process?
Can you specify which day and timeframe you would prefer for the interview? (Specify each option separated by a comma.)

We thank you for your time spent taking this survey.

Your response has been recorded.

2 Appendix B: Pilot Study Expert Interview Protocol & Script

Intro

– **Starting with briefly introducing the project and the purpose of the interview:**

  Script: "Thank you for agreeing to participate in this interview. My name is Samin Asnaashari, and I am conducting research as part of my Master's thesis in Human-Computer Interaction at Utrecht University. My thesis focuses on enhancing university admissions through advanced visualization and iterative design. The aim is to explore the current university admissions processes and identify areas where improvements can be suggested.

  Today, I would like to hear about your university admissions experiences and challenges and how you think technology, specifically interactive visualization tools, could improve decision-making.

  My overall strategy for my thesis is to comprehensively understand the admissions process on a global scale through current studies. Then, through these interviews, I aim to delve into the specifics of the admissions process at Utrecht University and compare it with best practices known across Europe and globally.

  Your insights will be valuable for my thesis and will help shape the development of a tool designed to support admissions committees."

– **Highlighting confidentiality and the use of information gathered during the interview by reviewing the consent form previously sent (see Appendix 1) and asking for permission to record or take notes.**

  Script: "Before we begin, I want to emphasize the confidentiality of this interview. Any information you provide will be used only for academic purposes and will be anonymized in my thesis. With your permission, I would like to record our conversation to ensure accuracy in gathering your insights. Do you have any questions before we start?"

Background Questions (Role and Experience in the Domain)

- **Q**: Can you describe your role and how long you have been involved in university admissions?

Goal: Understanding how involved the interviewee is in the admission process.

Domain Understanding - Problem Understanding

- **Q**: When was the last time you were involved in assessing an application?

- **Q**: Could you provide an overview of the steps you took from beginning to end?
- **Q**: Do you consistently follow the same process for every applicant, or do you find it varies? Can you explain any situations where you might deviate from the standard process?
- **Q**: Can you share some of the challenges you encounter during the admissions process? How do these challenges impact your ability to assess applicants?
- **Q**: Looking back, how confident do you usually feel about your admissions decisions? Can you recall any instances where you felt uncertain? How did you address this uncertainty?
- **Q**: In the context of admissions decisions, are there any rules (policy) or guidelines you are required to follow? Can you share the regulations and criteria you need to follow?
- **Q**: Did you find challenging or disagree with any of the regulations and criteria? How do these rules influence your decisions? Do you see anything missing?
- **Q**: Are more people usually involved in the process? Do you work with other officers? Are there any 4-eye principles? How are the applications allocated among the reviewers?
- **Q**: What tools or systems are currently used to manage and assess applications?

Goal: To understand the workflow and tools currently in place for highlighting areas for technological enhancement.

Requirement Elicitation

- **Q**: What data types are most critical in your admissions process and decision-making?
- **Q**: What data are missing for a proper fair decision outcome?
- **Q**: Do you believe that certain elements of the application could influence your evaluation of its other components?
- **Q**: Do you think the order in which the application is presented could impact your assessment?
- **Q**: If you had a 'magic wand' to create a tool or implement a change in the admissions process, what would it be? How would you envision this tool assisting you and your team?

Goal: To identify opportunities for introducing or improving data visualization tools in the process. Directly addresses the thesis goal of enhancing admissions through bias-aware visualization and design.

Final Thoughts & Comments

- **Q**: Are there any other experts, literature, or resources you recommend consulting to gain deeper insights into enhancing university admissions through visualization and design?

- **Q**: Is there anything else you think is important to consider in the context of enhancing university admissions through technology?

Closing

- **Thanking the participants for their time and insights, reminding them of the consent and how their input will be utilized.**
- **Outline the next steps in the research process.**

3 Appendix C: Feedback Session

Welcome to my HCI Master Thesis research study!

Visualization Evaluation Study

You are invited to participate in a **feedback study** conducted as part of a Master's Thesis research project in the field of **Human-Computer Interaction**, titled **"Visualizing Multi-Criteria Evaluations of Application Data in University Admissions: Supporting Holistic and Collaborative Decision-Making."**

Background and Purpose
The goal of this initial feedback session is to assess the clarity of the proposed data visualizations in supporting decision-making during university admissions, particularly in a collaborative context. This session seeks input on the visualizations before a full evaluation of the solution is conducted. Specifically, it aims to evaluate how well the visualizations aid the admissions process by providing a holistic view of applicants and enabling more confident, objective, and data-driven decisions. Additionally, the session will assess the overall readiness of the visualizations for final evaluation, focusing on clarity, and identifying any potential missing elements.

Study Point of Contact
This study is carried out by **Samin Asnaashari** (s.asnaashari@students.uu.nl) as part of her master's thesis in **Human-Computer Interaction** under the supervision of **Evanthia Dimara** (e.dimara@uu.nl) and **Christof van Nimwegen** (c.vanNimwegen@uu.nl).

Study Procedure

Overview
In this study, you will interact with "***EvaluationViz"*** tool design to support collaborative evaluation and decision-making in admissions. You will be asked to evaluate how clear and useful these visualizations are for the evaluation process and provide feedback through a questionnaire.

Tasks
- **Intro**: A short introduction to the tool, explaining its purpose and how to navigate.
- **Visualization Interaction**: You will first interact with several types of visualizations, including **Summary Chart, Tabular Panel Bar Chart, Stacked Bar Chart, and Lollipop Chart**. These visualizations represent input gathered from the evaluation form, focusing on different aspects of application data and corresponding metrics such as **score, weight, and uncertainty**. They provide various views of the collected inputs to offer a holistic and quantitative understanding of the application.
- **Questionnaire Feedback**: After exploring the visualizations, you will complete a questionnaire assessing the clarity, usability, and overall design of the visualizations. This feedback will help refine the tool to improve

transparency and reduce cognitive load in the admissions decision-making process.

Feedback Collection
Your feedback will be collected through this **questionnaire**, along with a **demographics questionnaire** for basic information for research purposes. All data is securely stored in this Qualtrics survey under the Utrecht University domain, and the results are only accessible to the researcher. The questionnaire collects only anonymous responses.

Your Rights
Participation is voluntary and based on your consent. You may withdraw at any time without any obligation to justify your decision. Upon withdrawal, all data collected until that point will be erased, excluding anonymized data already processed.

Study Approval
The Research Institute of Information and Computing Sciences has approved this study following an Ethics and Privacy Quick Scan. For complaints about the study conduct, contact **ics-ethics@uu.nl**. For personal data concerns or exercising GDPR rights, reach out to **privacy-beta@uu.nl**.
Visit **www.uu.nl/en/organisation/privacy** for more on privacy policies and your data rights.

By signing this form, I acknowledge and consent to the following terms for my participation in the research project titled "**Visualizing Multi-Criteria Evaluations of Application Data in University Admissions: Supporting Holistic and Collaborative Decision-Making**":

- **Age Confirmation:**
  - I confirm that I am 18 years of age or older.
- **Project Understanding:**
  - I confirm that the research project has been clearly explained to me. I have had the opportunity to ask questions and have received satisfactory answers. I understand that I can ask further questions at any time and that I have enough time to consider my participation.
- **Consent to Use of Material:**
  - I consent to the material I contribute being used to generate insights for the research project. The material will be anonymized and accessible only to the researcher.
- **Voluntary Participation and Withdrawal Rights:**
  - I understand my participation is voluntary, and I am free to withdraw at any time without needing to provide a reason. Upon withdrawal, any personal data collected from me will be erased.
- **Consent for Anonymized Data Use:**
  - I consent to the use of fully anonymized data in future publications and scholarly dissemination. I understand that anonymized data may be shared with other researchers for further research, and the University may publish it in data repositories for verification or broader access by researchers and other users. The data will be securely stored.

- **Data Deletion Request:**
  - I understand I have the right to request the deletion of any personal data collected from me.

○ I consent, begin the study

○ I do not consent, I do not wish to participate

The following questions focus on the **Clarity of the Visualizations**.

Clarity refers to how easily and accurately you can interpret the visualizations, ensuring the information is presented in a way that is understandable and free from confusion.

Q1: How would you describe the **clarity** of the **Summary Visualization** and its **interactive hover features** in helping you form a holistic view of the application?

Q2: How would you describe the **clarity** of the **Tabular Panel Bar Chart** and its **interactive hover features** in representing the evaluation data?

Q3: How would you describe the **clarity** of the **Stacked Bar Chart** and its **interactive hover features** in presenting the combined aspects of the application data?

Q4: How would you describe the **clarity** of the **Lollipop Chart** and its **interactive hover features** in presenting the metrics such as score, weight, and uncertainty per application data (Grades, Motivation Letter, ...)?

The following questions are designed to gather **overall impressions** of the visualizations.

The goal is to understand how the visualizations can support the evaluation and final holistic decision-making process and to identify any areas for improvement or missing elements.

Q5: Based on your interaction with the visualizations, which features do you think could be most helpful in supporting group decision-making, and why?

Q6: Were there any aspects of the visualizations that you found confusing or unnecessary? If so, please explain why.

Q7: What features or elements did you expect to see in the visualizations that were not present? How do you think these could have improved final decision-making process?

Q8: How would you describe the clarity of the **rubrics** and the **integration of application** data in the **evaluation form**? Do you feel the instructions are clear enough for anyone in an evaluation role to understand the tasks assigned to them?

Q9: How well do you think the **composite score** (which combines score, weight, and uncertainty into one number) will represent the overall evaluation of the applicant? Do you believe it will provide a trustworthy and accurate reflection of the applicant's suitability, or do you see any potential limitations? Please explain.

We thank you for your time spent taking this survey.

Your response has been recorded.

4 Appendix D: Phase 1 Evaluation - Without Visualization

Welcome to my HCI Master Thesis research study!
(First Phase)

You are invited to participate in the first phase of a research study conducted as part of a Master's Thesis project in the field of **Human-Computer Interaction (HCI)**.

This questionnaire consists of two parts: **Individual Input** and **After-Group Discussion Input**.

In this phase, you will review **two simulated applications**—**one individually and the other in a group setting**—for candidates applying to the **HCI master's program** for the **academic years 2022-2024**.

The application data will include components such as **Previous Education, Grades, Language Proficiency, CV, Motivation Letter, Reference Letters, and a Writing Sample**. You will be introduced to an **evaluation form**, along with a defined **simulated admissions policy**. For each data point, you will assess it based on the admissions policy by providing a **Score (your overall perceived rating of the data point), Weight (the significance of the data point in the decision-making process), and Uncertainty (your level of confidence in the assessment)**. Based on these assessments, you will make an **Accept/Reject** decision regarding the applicant's suitability—**first individually for the first applicant, and then collaboratively with your group for the second applicant to reach a final decision**.

Study Point of Contact
This study is carried out by **Samin Asnaashari** (s.asnaashari@students.uu.nl) as part of her master's thesis under the supervision of **Evanthia Dimara** (e.dimara@uu.nl) and **Christof van Nimwegen** (c.vanNimwegen@uu.nl).

Confidentiality Assurance
The data collected from this survey will be treated with the utmost confidentiality and will be anonymized to protect your privacy. All personal identifiers will be removed to ensure that the information cannot be traced back to any individual participant.

Your Rights
Participation is voluntary and based on your consent. You may withdraw at any time without any obligation to justify your decision. Upon withdrawal, all personal data collected until that point will be erased, excluding anonymized data already processed.

Study Approval
The Research Institute of Information and Computing Sciences has approved this study following an Ethics and Privacy Quick Scan. For complaints about the study conduct, contact ***ics-ethics@uu.nl***. For personal data concerns or exercising GDPR

rights, reach out to ***privacy-beta@uu.nl***. Visit ***www.uu.nl/en/organisation/privacy*** for more on privacy policies and your data rights.

By signing this form, I acknowledge and consent to the following terms for my participation in the research project:

- **Age Confirmation:**
  - I confirm that I am 18 years of age or older.
- **Project Understanding:**
  - I confirm that the research project has been clearly explained to me. I have had the opportunity to ask questions and have received satisfactory answers. I understand that I can ask further questions at any time and that I have enough time to consider my participation.
- **Consent to Use of Material:**
  - I consent to the material I contribute being used to generate insights for the research project. The material will be anonymized and accessible only to the researcher.
- **Voluntary Participation and Withdrawal Rights:**
  - I understand my participation is voluntary, and I am free to withdraw at any time without needing to provide a reason. Upon withdrawal, any personal data collected from me will be erased.
- **Consent for Anonymized Data Use:**
  - I consent to the use of fully anonymized data in future publications and scholarly dissemination. I understand that anonymized data may be shared with other researchers for further research, and the University may publish it in data repositories for verification or broader access by researchers and other users. The data will be securely stored.
- **Data Deletion Request:**
  - I understand I have the right to request the deletion of any personal data collected from me.

○ I consent, begin the study

○ I do not consent, I do not wish to participate

Questionaire after Individual Session:

Q1: What is your decision regarding the applicant’s suitability for the program based on the evaluation form—accept or reject?

Q2: In one sentence, can you describe how you came to your final decision?

Questionaire after Group Session:

Q3: What was the group's final decision regarding the applicant's suitability for the program—accept or reject?

Q4: In one sentence, can you describe how the group reached its final decision?

Q5: How confident are you that the group's evaluation fully considered all relevant inputs (both yours and others) and provided a clear understanding of the applicant's strengths, weaknesses, and fit for the program?
◯ Very confident – The group fully considered all inputs and has a clear understanding of the applicant's fit for the program.
◯ Somewhat confident – The group considered most inputs, but there are some uncertainties about the applicant's fit for the program.
◯ Not very confident – I feel the group missed some important inputs or is unclear about the applicant's fit for the program.
◯ Not confident at all – I am unsure about the group's evaluation and the applicant's fit for the program.

Q6: Did the group encounter any conflicting opinions during the discussion? How did the group identify, address, and ultimately resolve these differences?

We thank you for your time spent taking this survey.

Your response has been recorded.

5 Appendix E: Phase 2 Evaluation - Individual & Group

Welcome to my HCI Master Thesis research study!
(Second Phase)

You are invited to participate in the second phase of a research study conducted as part of a Master's Thesis project in the field of **Human-Computer Interaction (HCI)**.

This questionnaire consists of two parts: **Individual Input** and **After-Group Discussion Input**.

In this phase, you will review **two more simulated applications**—**one individually and the other in a group setting**—for candidates applying to the **HCI master's program** for the **academic years 2022-2024**.
The application data will include components such as **Previous Education, Grades, Language Proficiency, CV, Motivation Letter, Reference Letters, and a Writing Sample**. You will be introduced to an evaluation form and a simulated admissions policy. For each data point, you will assess it by providing a **Score (your overall perceived rating of the data point), Weight (the significance of the data point in the decision-making process), and Uncertainty (your level of confidence in the assessment)**.
Additionally, you will have access to several visualization tools, including a **Radar/Spider Summary Chart, a Tabular Panel Bar Chart, a Stacked Bar Chart, and a Lollipop Chart**. Based on these assessments and the visualizations, you will make an **Accept/Reject** decision regarding the applicant's suitability—**first individually for the first applicant, and then collaboratively with your group for the second applicant to reach a final decision**.

Study Point of Contact
This study is carried out by **Samin Asnaashari** (s.asnaashari@students.uu.nl) as part of her master's thesis under the supervision of **Evanthia Dimara** (e.dimara@uu.nl) and **Christof van Nimwegen** (c.vanNimwegen@uu.nl).

Confidentiality Assurance
The data collected from this survey will be treated with the utmost confidentiality and will be anonymized to protect your privacy. All personal identifiers will be removed to ensure that the information cannot be traced back to any individual participant.

Your Rights
Participation is voluntary and based on your consent. You may withdraw at any time without any obligation to justify your decision. Upon withdrawal, all personal data collected until that point will be erased, excluding anonymized data already processed.

Study Approval
The Research Institute of Information and Computing Sciences has approved this study following an Ethics and Privacy Quick Scan. For complaints about the study conduct, contact ***ics-ethics@uu.nl***. For personal data concerns or exercising GDPR

rights, reach out to ***privacy-beta@uu.nl***. Visit ***www.uu.nl/en/organisation/privacy*** for more on privacy policies and your data rights.

By signing this form, I acknowledge and consent to the following terms for my participation in the research project:

- **Age Confirmation:**
  - I confirm that I am 18 years of age or older.
- **Project Understanding:**
  - I confirm that the research project has been clearly explained to me. I have had the opportunity to ask questions and have received satisfactory answers. I understand that I can ask further questions at any time and that I have enough time to consider my participation.
- **Consent to Use of Material:**
  - I consent to the material I contribute being used to generate insights for the research project. The material will be anonymized and accessible only to the researcher.
- **Voluntary Participation and Withdrawal Rights:**
  - I understand my participation is voluntary, and I am free to withdraw at any time without needing to provide a reason. Upon withdrawal, any personal data collected from me will be erased.
- **Consent for Anonymized Data Use:**
  - I consent to the use of fully anonymized data in future publications and scholarly dissemination. I understand that anonymized data may be shared with other researchers for further research, and the University may publish it in data repositories for verification or broader access by researchers and other users. The data will be securely stored.
- **Data Deletion Request:**
  - I understand I have the right to request the deletion of any personal data collected from me.

○ I consent, begin the study

○ I do not consent, I do not wish to participate

Questionaire after Individual Session:

Q1: What is your decision regarding the applicant's suitability for the program based on the evaluation form—accept or reject?

Q2: In one sentence, can you describe how you used the visualizations (Radar/Spider Chart, Tabular Panel Bar Chart, Stacked Bar Chart, Lollipop Chart) to help inform your individual decision?

Questionaire after Group Session:

Q3: What was the group's final decision regarding the applicant's suitability for the program—accept or reject?

Q4: Describe in one sentence how your group used the visualizations (Radar/Spider Chart, Tabular Panel Bar Chart, Stacked Bar Chart, Lollipop Chart) to support the evaluation and reach a final decision.

Q5: Please rank the visualizations (Radar/Spider Chart, Tabular Panel Bar Chart, Stacked Bar Chart, Lollipop Chart) from most to least helpful in supporting the group's evaluation and helping form a final decision.

Radar/Spider Chart	1
Tabular Panel Bar Chart	2
Stacked Bar Chart	3
Lollipop Chart	4

Q6: In one sentence, how would you describe your group's experience in forming a holistic view of the applicants using visualizations compared to your previous experience without visualizations?

Q7: How effective do you think the visualizations would be in helping your group recall and understand the reasoning behind the collaborative decision made about the applicant's compatibility with the program if you returned to it later?

◯ Very effective – The visualizations would clearly reflect the group's reasoning behind the decision.

◯ Somewhat effective – The visualizations would help, but some details of the group's reasoning might be harder to recall.

◯ Not very effective – The visualizations would offer limited help in recalling the group's reasoning behind the decision.

◯ Not effective at all – The visualizations would not help the group understand or recall the reasoning behind the decision.

We thank you for your time spent taking this survey.

Your response has been recorded.

6 Appendix F: Demographics Questionnaire

Thank you for your participation in my HCI Master's Thesis research study!

This questionnaire collects basic demographic information for research purposes, and all responses will remain anonymous.

Q1: What is your gender identity?
○ Male
○ Female
○ Non-binary
○ Other
○ Prefer not to say

Q2: What is the age range to which you belong?
○ 20-29 years old
○ 30-39 years old
○ 40-49 years old
○ 50-59 years old
○ 60-69 years old

Q3: What is your nationality? [ ]

Q4: What is the highest level of education you have completed?
○ High School or equivalent
○ Associate Degree
○ Bachelor's Degree
○ Master's Degree
○ Doctorate (PhD)
○ Other

Q5: What is your current professional title and role? [ ]

Q6: How familiar are you with the practice of information visualization, which involves representing data in meaningful, visual formats such as charts and dashboards to facilitate comprehension and decision-making?

○ Not Familiar: I have no prior knowledge or experience with information visualization
○ Slightly Familiar: I have heard of information visualization but have limited understanding or experience with it
○ Moderately Familiar: I have some knowledge and experience with information visualization and have used basic tools and techniques

◯ Very Familiar: I have a good understanding of information visualization and regularly use various tools and techniques to create visual representations of data
◯ Extremely Familiar: I am highly proficient in information visualization, with extensive experience in creating complex visualizations and dashboards for data analysis and decision-making

Q7: In the context of evaluating applicants for their suitability for positions like internships, external or internal job roles, or educational admissions, which of the following statements most accurately reflects your level of experience and familiarity?
◯ No Experience: I have never been involved in assessing applicants for any of these contexts
◯ Limited Experience: I have been involved in assessing applicants on a few occasions but have limited familiarity with the process
◯ Moderate Experience: I regularly assess applicants and am familiar with the general process, but I am not deeply involved in decision-making
◯ Extensive Experience: I frequently assess applicants and am actively involved in decision-making
◯ Expert Level Experience: I am highly experienced in assessing applicants and have a deep understanding of the evaluation criteria and decision-making process

Q8: In one sentence, describe your experience evaluating applications. Have you typically done this individually or collaboratively? How do you form a final accept or reject decision?

Q9: In one sentence, describe your experience using visualizations in the decision-making process for hiring or educational evaluations. What types of visualizations did you use, and how did they aid your decision-making?

We thank you for your time spent taking this survey.

Your response has been recorded.

7 Appendix G: Simulated Application Data & Rubric

Below are the two simulated application datasets and the corresponding simulated rubric policy used for evaluation:

Application 1:

Application for HCI Master's Program (2022-2024)

Previous Education:

- **University:** Southern New Hampshire University (SNHU)
- **Degree:** B.Sc in Data Science with a Minor in Artificial Intelligence
- **GPA:** 3.1 (on a 4.0 scale)

Transcript:

- Introduction to Artificial Intelligence: 7/10
- Data Structures and Algorithms: 7/10
- Machine Learning: 6/10
- Natural Language Processing: 8/10

Language Proficiency:

- **English:** Native Speaker
- **Spanish:** B2 level, DELE certificate

CV:
Graduated from Southern New Hampshire University (SNHU) with a B.Sc in Data Science. Worked as a research assistant at the SNHU Department of Computing, where I contributed to projects focused on the ethical use of AI in decision-making systems. Interned as a junior data scientist at a regional tech firm, specializing in developing predictive models to improve user engagement for mobile applications. I am particularly interested in how artificial intelligence can be used to enhance human-computer interaction, and I aim to explore this further through the HCI Master's program.

Motivation Letter:
I am writing to express my strong interest in pursuing the Human-Computer Interaction (HCI) Master's program at Utrecht University. During my undergraduate studies at Southern New Hampshire University, I discovered my passion for artificial intelligence and its potential to transform the way humans interact with technology. In particular, I was drawn to projects where AI was used to create personalized user experiences, which led to my work on developing machine learning models aimed at improving user engagement. The HCI Master's program at Utrecht, with its emphasis on user-centered design and personalized computing, aligns perfectly with my goal to further explore the relationship between AI and human interaction. I am confident that the program will provide me with

the academic and practical foundation necessary to contribute to this evolving field.

Reference Letters:
Prof. Dr. Emily Greene, Professor of Data Science at Southern New Hampshire University: "As a professor in the Data Science department, I had the pleasure of supervising [Applicant Name] during their undergraduate research assistantship. Their work on AI-driven systems demonstrated a unique ability to approach complex problems with innovative solutions. In particular, their contribution to our project on ethical AI in decision-making systems showcased both their technical skills and their commitment to addressing the broader implications of AI. [Applicant Name] displayed great initiative and an impressive ability to bridge technical knowledge with user-centered applications. I strongly believe that they will excel in the Human-Computer Interaction Master's program at Utrecht University and make meaningful contributions to the field."

Writing Sample:
Submitted a research paper titled "Leveraging AI for Ethical Decision-Making in User Interfaces: A Framework for Fairness in HCI Systems". This paper explores how artificial intelligence can be designed to ensure fairness in human-computer interaction, particularly in systems that make decisions on behalf of users. The research focuses on the development of an AI framework that balances personalization with ethical considerations, ensuring that user experiences are enhanced without bias. The methodology outlines the creation of machine learning models trained to recognize potential biases in user interface design, paired with real-time feedback loops that adapt to user needs while maintaining fairness across all interactions. The conclusion highlights the importance of integrating ethical AI into HCI systems to foster trust and transparency in technology-driven interactions.

Application 2:

Application for HCI Master's Program (2022-2024)

Previous Education:

- **University:** Fontys University of Applied Sciences
- **Degree:** B.Sc in Information Communication Technology (ICT)
- **GPA:** 3.6 (on a 4.0 scale)

Transcript:

- Operating Systems: 9/10
- C# Programming: 8/10
- Statistics: 6/10
- TTIL (ITIL Foundations): 5/10
- Database Management: 7/10

Language Proficiency:

- **English:** C2 level, IELTS score of 7.0
- **German:** A2 level, Goethe-Zertifikat, Pass

CV:
Graduated with a B.Sc in ICT from Fontys University, where the applicant also pursued a pre-master in Computer Engineering at the Technical University of Eindhoven (TU/e). The applicant has experience working as a lead front-end engineer at a tech company in Amsterdam, specializing in UX/UI design and project management.

Motivation Letter:
With this letter, I express my strong interest in applying for the Human-Computer Interaction Master's program at Utrecht University. My passion for HCI stems from my background in both technology and user experience design. During my time at Fontys, I was captivated by how the visual design of interfaces can influence user patience and decision-making. My professional journey as a front-end engineer, combined with my strong foundation in software engineering and UX/UI design, has only deepened my interest in the intersection between technology and human behavior. Utrecht's focus on empathic and personalized computing perfectly aligns with my desire to create impactful user interfaces that enhance accessibility and user satisfaction.

Reference Letters:
Prof. Dr. Jan de Vries, Professor of Information Technology, Fontys University of Applied Sciences: Lena van der Meer was one of the most promising students I had the privilege to teach during my tenure. She consistently demonstrated a unique ability to blend creative design thinking with technical proficiency. Her work on human-computer interaction, particularly her final project focusing on enhancing user experience for aging populations, showcased her commitment to creating accessible, user-centered designs. Lena's academic rigor, combined with her practical experience as a student assistant and later as a lead engineer, makes her an ideal candidate for the HCI program at Utrecht University. I am confident that she will excel in your program and contribute meaningfully to the field of HCI.

Writing Sample:
This paper, titled 'Image Analysis and Aging: Improving Predictive Accuracy in Medical Imaging Through UX and AI Integration,' explores the application of machine learning techniques to analyze age-related changes in facial features. The project aimed to enhance age detection algorithms used in medical imaging, with a focus on improving both the accuracy of predictions and the user interface for healthcare professionals. The abstract outlines the integration of deep learning models with user-centered design principles, ensuring that medical practitioners can easily navigate complex diagnostic tools without requiring advanced technical knowledge. In the introduction, the problem of age-related visual changes in diagnostic imaging is presented as a critical challenge for aging populations, and the need for accurate, reliable tools in healthcare settings is emphasized. The methodology section details the creation of a deep learning model trained on a dataset of facial images, paired with a user interface that simplifies the presentation of complex data. By prioritizing transparency and usability, the study bridges the gap between technical complexity and practical application in healthcare environments. The conclusion emphasizes the importance of human-computer interaction in medical technology, showing

that improved user interface design not only enhances usability but also increases trust in AI-driven tools among non-technical users. This work demonstrates the potential for UX design to significantly impact the adoption of advanced diagnostic systems in medical fields, particularly for age-related health issues.

Simulated Rubric:

- **Previous Education:**
  - **Degree Requirement:** Bachelor's degree in Computer Science, Information Technology, Information Science, Psychology, or a related field. For students from universities of applied sciences (HBO), completion of a pre-master's program may be required.
  - **GPA:** A minimum GPA of 3.0 or equivalent on a relevant grading scale. For international students, GPA equivalence will be assessed based on the country-specific grading system.
  - **Course Relevance:** Completion of coursework related to HCI, user experience design, software engineering, data analysis, or human-centered AI. International students should demonstrate equivalency in their local education systems.
- **Grades Transcript:** Review academic transcripts for strong performance in courses related to HCI, user interface design, programming, data analysis, and psychology. For students from universities of applied sciences (HBO), relevant work experience and practical projects should also be considered. International transcripts should be accompanied by certified translations if not in English.
- **Language Proficiency:**
  - **English:** Proof of English proficiency (IELTS minimum 6.5, or TOEFL equivalent: 93 iBT). For students from non-English speaking countries, additional proof of language competence may be required.
  - **Certificates:** Language proficiency certificates must be valid and obtained within the last two years. Students from international institutions where English is the primary language of instruction may be exempt.
- **CV:** Relevant work or research experience in HCI, UX design, artificial intelligence, or related fields is strongly recommended. For students from HBO programs, practical experience or internships should weigh significantly.
- **Motivation Letter:** A clear motivation statement detailing the candidate's interest in HCI, their research ambitions, and how the program aligns with their career goals. International students should also describe any cross-cultural experiences or perspectives that enrich their interest in HCI.
- **Reference Letters:** Strong academic or professional reference letters, ideally from those familiar with the candidate's work in HCI, UX, or related research. For international students, references should ideally come from institutions or companies that are internationally recognized.

- **Writing Sample:** A relevant writing or research sample that demonstrates the candidate's analytical abilities, research potential, and presentation of information in the HCI domain. For HBO graduates, portfolio projects and practical reports may also be considered alongside academic writing. International students should ensure any non-English samples are translated.