Class-Based vs. Instance-Based Interface Design for Unanticipated User-Generated Content in Student Life Reporting

by

Maryam Mofrad

A thesis submitted to the School of Graduate Studies in

partial fulfillment of the requirements for the degree of

Master of Science in Management.

Department of Business Administration

Memorial University

May 2025

St. John's, Newfoundland and Labrador, Canada

Abstract

This thesis investigates the effectiveness of Instance-Based versus Class-Based interface designs in capturing and managing user-generated content within a student life context. With the proliferation of digital platforms, the volume and variety of user-generated content have surged, challenging traditional structured user interface designs. Traditional Class-Based interface designs often fail to accommodate the dynamic nature of user-generated content, leading to the potential loss of valuable insights. In contrast, Instance-Based interface designs offer a flexible, potentially improving data representation and usability. This thesis explores the consequences of using Class-based versus Instance-Based interface design to collect user-generated content, focusing on student life reporting. The study is driven by questions on how these two interface configurations compare in their capacity to manage the diverse nature of user-generated content. By applying both designs in a real-world setting and analyzing the resultant data, the study aims to furnish empirical insights into the suitability of each design for user-generated content data collection and management. The findings suggest that while Instance-Based interface design offers significant improvements in flexibility and data representation, it also poses challenges in terms of complexity and user engagement. This research contributes to the broader discourse on data models in the context of big data, highlighting the potential of Instance-Based interface design to enhance the collection of user-generated content.

Keywords: Conceptual modeling, interface design, data models, instance-based, class-based, student life reporting.

Acknowledgments

I want to express my sincere gratitude to all those who have contributed to completing this master's thesis.

First and foremost, I am deeply grateful to my advisor, Dr. Jeffrey Parsons, for his invaluable guidance, mentorship, and support throughout this research journey. His expertise and dedication have been instrumental in shaping this thesis and enhancing my research skills.

A special mention goes to my fellow classmates, who provided continuous support, shared valuable knowledge, and offered insightful discussions. Your contributions were a cornerstone of my academic experience, and I am thankful for the camaraderie and collaboration we shared.

I am incredibly grateful to the participants of this study for generously sharing their time and insights. Their willingness to be part of this research significantly enriched the depth and quality of the findings.

Table of contents

List of Figures

List of Tables

List of abbreviations

1. ACM - Association for Computing Machinery
2. AI - Artificial Intelligence
3. CAD/CAM - Computer-Aided Design/Computer-Aided Manufacturing
4. ERD - Entity-Relationship Diagram
5. GIS - Geographic Information Systems
6. IBDM - Instance-Based Database Model
7. IBID - Instance-Based Interface Design
8. IQ - Information Quality
9. MUN - Memorial University of Newfoundland
10. NoSQL - Not Only SQL
11. OLTP - Online Transaction Processing
12. QL - Query Language
13. RCT - Randomized Controlled Trial
14. RDM - Relational Database Model
15. CBID – Class-Based Interface Design
16. SQL - Structured Query Language
17. UGC - User-Generated Content
18. VBA - Visual Basic for Applications

Chapter 1

1. Introduction

1.1. Background

User-generated content (UGC) refers to various forms of media and information created and shared by users rather than professionals. This type of content often emerges in dynamic and diverse contexts, making the Internet more interactive and participatory (Tomaiuolo, 2012). The rise of UGC across multiple platforms has created challenges and opportunities for data collection and analysis. Traditional data models, particularly class-based models, have long been used to structure and analyze data within predefined categories. In class-based models, data is organized into predefined classes or categories, each representing a specific type of entity or phenomenon (e.g., product types, species, or other categories) with a set of predefined properties to which specific instances must conform (Lukyanenko, 2014). While these models have been foundational, their rigidity often limits their capacity to handle the dynamic and unpredictable nature of UGC, potentially leading to a loss of unanticipated insights and information (Lukyanenko, 2014; Parsons & Wand, 2000).

The inherent limitations of class-based models, such as their inability to evolve or adapt to new data requirements and the constraints they impose on capturing diverse data types, necessitate a shift in data modeling approaches to accommodate the rapidly changing digital landscape. This has led to the development and exploration of the instance-based data model (IBDM). The IBDM prioritizes flexibility and granularity by focusing on individual instances rather than predefined classes (Parsons & Wand, 2000; Lukyanenko et al., 2019). This approach is crucial for effectively capturing the inherent variability and unpredictability of UGC, enabling a more accurate and nuanced representation of real-world phenomena. By emphasizing adaptability, the IBDM allows for the exploration and analysis of unstructured data and accommodates new data types without significant reconfiguration.

Empirical studies have investigated the practical implications of using instance-based versus class-based models, particularly concerning their impact on data collection and the ability to capture unanticipated insights. These studies use a variety of methodologies, including case studies, surveys, and experiments, to explore how each model influences user performance, flexibility in data representation, and overall usability (Saghafi & Wand, 2014; Recker et al., 2019). Findings indicate that while the IBDM improves the ability to discover patterns and retrieve information effectively, challenges such as defining attribute hierarchies and ensuring the model's relevance within specific domains persist.

1.2. Objectives of the Thesis

Building on the discussion of data models in the background, this thesis applies them to interface design, comparing Class-Based and Instance-Based Interface Designs for managing UGC, with a specific focus on student life reporting. The objective of this thesis is to analyze the limitations of Class-Based interface design (CBID) in handling UGC and investigate the theoretical foundations and practical applications of Instance-Based interface design (IBID). It designs and implements a comparative study to evaluate the performance of IBID and CBID in capturing UGC, highlighting both the advantages and challenges of using IBID in real-world applications. Through this research, the thesis aims to demonstrate the flexibility, accuracy, and ability of IBIDs to capture unanticipated data, contributing valuable insights to for the design of data collection systems. By applying both designs in a practical setting and analyzing the resultant data, this research aims to provide empirical insights into the effectiveness of each design for managing UGC, with the objective of contributing to the broader discourse on interface designs in the context of dynamic and complex data environments.

1.3. Structure of the Thesis

The structure of this thesis is as follows. The Introduction outlines the research problem, its significance, and objectives. Chapter 2 explores relevant literature on data models, focusing on the evolution from relational database models (RDM) to IBDM and their implications in interface design. Chapter 3 presents the research hypotheses grounded in the limitations of CBID and the potential benefits of IBDM. Chapter 4 details the randomized controlled trial (RCT) methodology, including survey flow and participant recruitment. Chapter 5 analyzes the collected data, comparing the effectiveness of Class-Based and Instance-Based models. Unlike previous research, this analysis uses qualitative methods, including thematic and sentiment analysis.

Chapter 6 addresses the study's constraints and suggests areas for further exploration. Finally, Chapter 7 summarizes the findings and their implications for database design and management, emphasizing IBDM's potential to enhance data quality and user engagement in diverse data environments.

1.4. Significance and Contribution of the Study

This study advances the understanding of how interface design influences UGC collection, particularly in student life reporting. While prior research has explored structured and unstructured data collection in citizen science and information systems, this study applies those insights to a novel domain, demonstrating how IBID enhances data completeness and authenticity in qualitative feedback environments.

The findings of this research will benefit several stakeholders, including UX designers, survey methodologists, and policymakers involved in large-scale data collection and analysis. By highlighting the trade-offs between CBID and IBID, this work provides actionable recommendations for designing more effective and user-friendly reporting interfaces.

It identifies key challenges of IBID, such as increased cognitive load and survey abandonment rates, while proposing hybrid models that balance structured prompts with open-ended responses.

Empirically, this study underscores the potential of IBID to enrich data collection by capturing nuanced and detailed information, offering practical guidance for researchers and practitioners. It extends the discourse on interface design's role in user engagement, paving the way for future research in diverse contexts such as healthcare, marketing, and personalized services.

By highlighting the strengths and limitations of IBID and CBID, this thesis offers valuable insights into optimizing interface design for UGC collection. The findings support the broader adoption of instance-based approaches in dynamic data environments, enhancing both research and practical applications in user-driven content management.

Chapter 2

2. Data Models and Interface Design

Databases are structured data collections representing specific segments of the real world, capturing essential details about relevant phenomena. The data model within a database provides the structural framework, outlining the abstract organization of data to reflect real-world segments accurately. A data model is essentially a blueprint that defines how data elements are organized, categorized, and interrelated within a system. It provides a framework that governs the static properties of data (such as objects and their relationships) as well as dynamic behaviors, such as operations that can be performed on the data (Silberschatz et al., 1996). This foundational structure impacts not only the backend organization of databases but also how users interact with the data through interfaces.

Interface design, in turn, focuses on how users interact with a system, which includes the visual elements, navigation flows, and interactive mechanisms that allow users to input, access, and manipulate data. Interface design encompasses the creation of user-friendly and efficient ways for individuals to interact with data systems. It includes the layout of forms, the use of interactive elements like buttons and drop-down menus, and the visual representation of data relationships (Kurosu, 2020). Data models and interface design are intrinsically linked, as the structure and flexibility of the data model determine how intuitive and adaptable the interface can be. The relationship between data models and interface design is significant because the structure of the data model directly influences how information is presented and accessed in the user interface (Brodie, 1984). Effective interface design depends on the underlying data model's ability to support clear, intuitive user interactions, making the user experience a key outcome of database design decisions.

2.1. Developments in database technologies

The landscape of database design has undergone significant transformations since its inception in the mid-20th century. These changes reflect the evolving needs of businesses and technological advancements, leading to various generations of database systems, each characterized by distinct models and capabilities. The evolution from the first generation of databases, which primarily focused on efficient data storage and retrieval, to the more sophisticated systems of today demonstrates a shift towards increased flexibility, scalability, and performance. The following Table 2.1 details the chronological development of database designs. A summary of these developments through time is presented in Figure 2.1.

Table 2.1 Evolution of Database Systems by Generation and Key Developments

Generation	Description	References
First Generation (mid-1960s and early 1970s)	Development of general-purpose database management systems Based on hierarchical and network data models.	Tsichritzis and Lochovsky (1976); DBTG Codasyl (1971); Taylor and Frank (1976)
	Exemplified by systems like IDS (network), IMS (hierarchical), IDMS (network)	Meltz et al. (2004)

	High performance and throughput through low-level procedural operations for navigating linked records	
Second Generation (early 1970s)	Based on the relational data model (RDM), representing data as tuples in relations offers substantial data independence.	Codd (1970)
	Led to the adoption of set-oriented declarative query languages, with SQL (Chamberlin and Boyce, 1974) becoming the standard.	Chamberlin and Boyce (1974)
Early 1980s (Complex Database Applications)	Issues with Relational Data Models (RDM) for Computer-Aided Design/Computer-Aided Manufacturing (CAD/CAM), Geographic Information Systems (GIS), and applications with complex data structures.	Maier (1989); Zaniolo et al. (1985)
	Emergence of object-oriented database systems: data stored as true objects identified by OIDs	Beeri (1990); Kim (1990)
	Emergence of object-relational database systems, incorporating object-oriented features into relational systems. Major systems like Oracle and DB2 adopt	Stonebraker and Moore (1995)

	extensions, mapping, and joins done automatically.	
Early 2000s (Advances in Web Technology)	Explosion of structured, semi-structured, and unstructured data from the web, social networking, mobile devices, IoT	
	Challenges in achieving these requirements through traditional relational database systems	McAfee et al. (2012); Gray et al. (1996); Helland (2007)
Fourth Generation (early 2000s onward)	Emergence of NoSQL stores to handle high availability and scalability requirements of global-scope applications	Abadi (2012); Brewer (2000); Stonebraker (2010b); Cattell (2011)
	NoSQL features: flexible data models, weak consistency transaction models, use of distributed indices, hashing, and caching.	Grolinger et al., 2013; Araújo et al., 2023
	Not a replacement but a remedy for relational systems' shortcomings in handling big data	
Fifth Generation (late 2000s)	Emergence of NewSQL data stores to tackle scalability and reliability requirements of modern OLTP applications	Stonebraker (2012)
	New architecture for improved scalability and performance, maintaining some relational	

	functionalities like multikey ACID transactions	

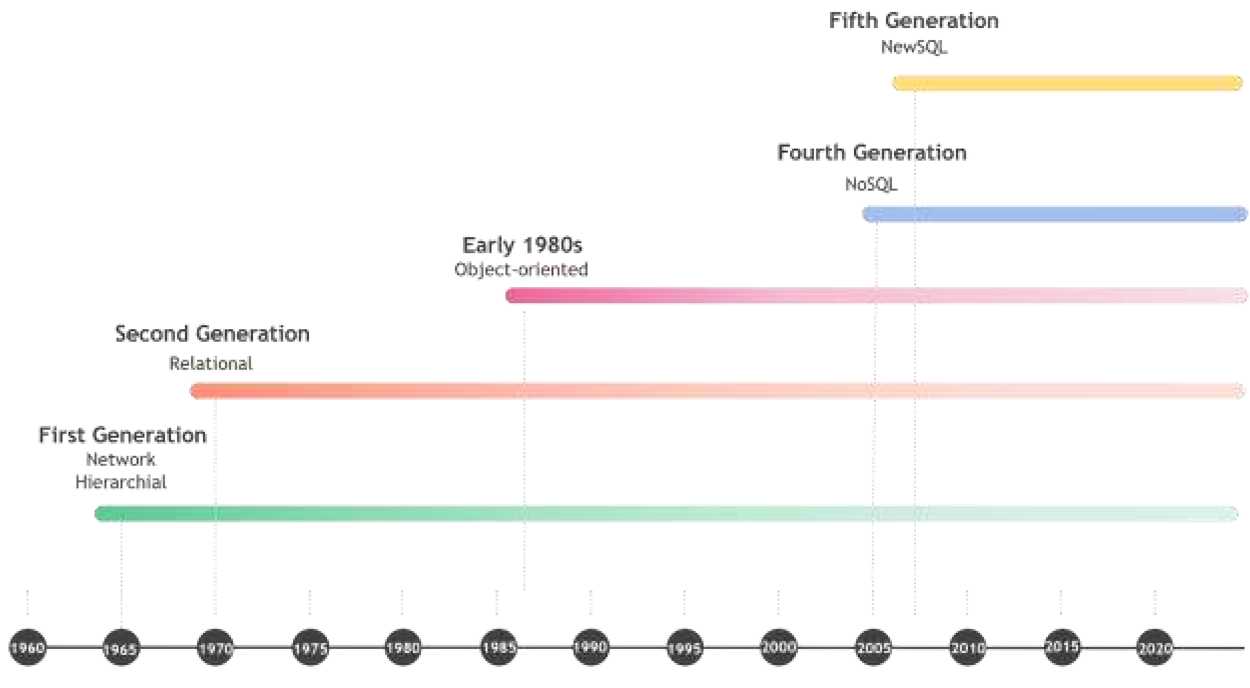

Figure 2.1 Evolution of major database technologies.

2.2. Class-Based Data Models

The Entity-Relationship Model (ERM) is a fundamental technique in database design that was introduced by Chen (1976). Chen's work emphasized the natural representation of the real world as entities and relationships, making the ER model an intuitive and effective approach to database design (Watt, 2014).

At the conceptual level, the ER model defines key concepts such as entities, relationships, attributes, and cardinalities (Parsons & Wand, 2000). These elements work together to create a comprehensive representation of the data and its interactions. The ER model's capacity to clearly structure and define data ensures its ongoing relevance, as Verdonck et al. (2018) confirm, making it a principal model for conceptual database design.

2.2.1. Advantages of Class-Based Data Model

Class-based data models offer several distinct advantages, primarily in their ability to provide consistency, uniformity, and structure. By enforcing a standardized format, these models simplify data analysis and minimize the need for extensive post-processing, enhancing data consistency and reducing errors (Lukyanenko et al., 2017). This structured approach is particularly useful when contributors are familiar with the domain, as it minimizes training requirements and increases the accuracy of data collection.

Additionally, class-based models align well with projects that have well-defined goals and data parameters. They ensure that data remains within expected boundaries, supporting efficient and targeted analysis (Shanks et al., 2008). By focusing on common attributes, these models simplify complex real-world scenarios, resonating with human cognitive processes. This alignment reduces cognitive load, improves communication, and enhances understanding (Shanks et al., 2008).

Furthermore, these models support inferential reasoning by allowing contributors to make informed guesses based on predefined attributes, thus enhancing data organization and interoperability (Parsons & Wand, 1997). Their structured nature promotes stability and clarity in data management, allowing for efficient organization, retrieval, and access control, thereby strengthening data security (Shanks et al., 2008).

2.2.2. Limitation of Class-Based Data Model

Despite its advantages, Class-Based data models also have several limitations, particularly in terms of flexibility, context biases, and the representation of unique information. One major drawback is the system's rigidity, which demands stable categories and boundaries. This rigidity makes it difficult to capture evolving or unique data, often necessitating costly updates to database structures, programming, and interfaces when schema changes are required (Lukyanenko et al., 2017). In dynamic domains where data types frequently emerge, this inflexibility becomes a significant barrier.

Additionally, the focus on predefined attributes in Class-Based systems can introduce context-setting biases. Classes may guide users toward familiar features, limiting their ability to identify new or unexpected phenomena (Lukyanenko et al., 2018). This constraint may reduce creativity and engagement, as contributors are less likely to explore beyond the established categories.

Moreover, the reliance on well-defined categories poses challenges when data does not fit neatly into the existing structure. Inaccuracies may arise, particularly in fields where new data continuously evolves, making it challenging to adapt the schema accordingly (Lukyanenko et al., 2018; Shanks et al., 2008). This issue is further compounded when the data model must accommodate evolving requirements or capture unexpected information.

The need for extensive training also presents a barrier, particularly for non-experts or amateur contributors who may struggle with understanding and applying established categories (Lukyanenko et al., 2017).

2.3. Instance-Based Data Model (IBDM)

The IBDM represents a significant departure from traditional Class-Based models by emphasizing the unique characteristics of individual data instances rather than grouping them into predefined classes. Traditional models, such as relational and object-oriented models, often overlook the distinct attributes of each instance, focusing instead on shared characteristics for generalization. This can limit the depth of analysis and the applicability of data, particularly in personalized fields like medicine or customer behavior analysis.

The IBDM focuses on capturing and utilizing the distinct properties of each data instance, allowing for more precise and tailored analysis. It provides a more flexible and detailed representation of data, aligning closely with modern, complex data environments characterized by high variability (Lukyanenko et al., 2018; Parsons & Wand, 2000; Parsons & Su, 2006). By focusing on individual instances, the IBDM accommodates the nuances of data without the constraints of predefined categories, offering a more accurate representation (Lukyanenko et al., 2018).

Furthermore, the IBDM enhances data integrity and relevance by maintaining individuality, which is crucial in dynamic environments where data constantly evolves. By not relying on predefined categories, these models offer flexibility and adaptability, accommodating new or unexpected data types effectively. These models support environments requiring high customization, as seen in big data and UGC, where data does not fit neatly into set categories (Lukyanenko et al., 2018; Parsons & Su, 2006).

2.3.1. Advantages of IBDM

The IBDM offers a transformative approach to handling data in various database management systems, particularly in environments where flexibility, accuracy, and efficiency are paramount. This comprehensive model facilitates more nuanced and effective data management practices by emphasizing individual instances over rigid, predefined classifications, offering significant advantages in multiple fields, from cloud storage solutions to user-generated data management (Araújo et al., 2023; Parsons & Wand, 2000; Parsons & Su, 2010; Lukyanenko, 2014; Lukyanenko et al., 2019).

Enhanced Flexibility and Adaptability

The IBDM provides unparalleled flexibility in data management by allowing each data instance to be treated independently, devoid of strict adherence to a predefined class structure. This flexibility proves particularly advantageous in managing changes in data structures and relationships, as it eliminates the need to overhaul entire schemas when modifications are required (Parsons & Wand, 2000). Such an approach not only simplifies the integration of schemas from different databases but also significantly reduces conflicts and complexities typically associated with schema integration. It thereby supports easier schema evolution, reduces the operational overhead, and diminishes the complexities tied to traditional Class-Based models (Parsons & Wand, 2000).

Moreover, by not confining data to rigid class structures, the IBDM reduces the risk of information loss, capturing a broader spectrum of data and reflecting the nuances of UGC more accurately (Lukyanenko, 2014). This method is particularly beneficial in environments where data properties might not be initially known or are subject to change, allowing properties to be defined or redefined as needed (Parsons & Su, 2010).

Improved Data Accuracy and Security

The IBDM enhances data accuracy by allowing more detailed and specific information to be recorded for each instance, which reduces the generalization errors that Class-Based models might introduce (Lukyanenko, 2014).

Discovery of New Insights and Real-World Applications

The Instance-Based approach is conducive to discovering new, unexpected data because it does not restrict contributors to predefined classes or categories. Contributors can report observations as they see them, which can lead to new insights and findings, particularly useful in fields such as crowdsourcing and UGC (Lukyanenko et al., 2019). Moreover, IBDMs are particularly effective in contexts where data can be highly variable and where contributors have different levels of expertise and familiarity with the data being collected. This method accommodates a wider variety of data inputs and structures, reflecting real-world complexities more effectively (Lukyanenko et al., 2019).

2.3.2. Limitation of IBDM

While the IBDM offers numerous advantages in terms of flexibility, detail, and adaptability, it also introduces several challenges that can complicate data management and system operation. These challenges primarily stem from the model's inherent complexity, its demand for sophisticated management, and the potential trade-offs concerning data consistency and system performance.

Adaptation and Learning Curve

The shift from Class-Based to Instance-Based modeling requires a significant learning curve for database designers and administrators. Existing systems and processes might also need considerable adaptation to fit this new paradigm, complicating the transition and potentially increasing the time to proficiency for the personnel involved (Parsons & Wand, 2000). The Instance-Based model might also dilute clear hierarchies and inheritances found in Class-Based systems, potentially leading to confusion or mismanagement of data relationships (Parsons & Wand, 2000).

Performance and Consistency Trade-offs

IBDMs often use eventual consistency to improve performance, which may lead to inconsistencies where not all nodes are updated simultaneously. This can result in accessing outdated data, which might not be acceptable for applications requiring strong consistency (Araújo et al., 2023). Managing different consistency levels to balance between availability and consistency can be complex, requiring sophisticated techniques to assess the impact of different design choices on system behavior and performance. While these models generally improve performance, certain configurations and consistency levels might increase latency, particularly when strong consistency is required or during high system load scenarios (Araújo et al., 2023).

2.4. The Influence of Data Models on Interface Design

Data models influence interface design by determining the types of inputs users can provide and the pathways they can navigate within a system. A well-designed data model should not only organize data efficiently but also support clear and intuitive user interactions. The design of the user interface should thus reflect and enhance the structure set by the data model, allowing users to interact with the system efficiently. For example, a class-based data model, which relies on predefined categories and relationships, often results in interfaces that guide users through a fixed set of steps and interactions. This structured approach works well when data entries are uniform and predictable, ensuring that interfaces remain consistent and reliable over time (Silberschatz et al., 1996). However, such rigidity can be a limitation when dynamic or unstructured data must be accommodated, as class-based models may not be able to adapt to new data types or structures without significant reconfiguration (Brodie, 1984; Lukyanenko et al., 2017).

In contrast, instance-based models, which prioritize flexibility and adaptability, allow for more dynamic interfaces that accommodate various user inputs. These models are designed to capture the unique characteristics of individual data instances rather than fitting all data into predefined categories (Lukyanenko et al., 2018). This flexibility enables interfaces to be more responsive, supporting real-time adjustments and evolving data inputs. Such adaptability is particularly beneficial in modern applications like UGC platforms, where the diversity and unpredictability of data make rigid schemas impractical (Pedersen et al., 2018). The flexibility offered by instance-based models leads to the development of interfaces that can dynamically evolve to accommodate new patterns and types of information (Abrahão et al., 2021).

The literature on survey interface design emphasizes that enhancing user engagement is critical for improving the quality of data collected. Abrahão et al. (2021) discuss how user interface adaptation can significantly improve user experience by customizing the interface to meet the specific needs and preferences of users. In terms of unstructured data, interfaces designed to handle such input often incorporate flexible fields that allow users to express themselves freely. Such designs can lead to higher levels of user engagement, as users feel less constrained and more inclined to provide detailed and varied responses. However, the trade-off is the potential for inconsistency in data quality, as unstructured responses may vary significantly in detail and clarity (Kurosu, 2020).

The design of survey interfaces, whether class-based or instance-based, directly impacts user engagement and the overall quality of data collected. Research indicates that well-designed interfaces that minimize cognitive load and provide clear, intuitive pathways for users tend to yield higher-quality data (Lukyanenko et al., 2017; Pedersen et al., 2018). For class-based interfaces, maintaining a balance between providing enough structure to guide users and avoiding excessive rigidity is essential. Overly rigid designs can deter users, leading to lower engagement levels and potentially reduced data accuracy.

For instance-based interfaces, the challenge lies in managing the variability of input while ensuring that the collected data remains relevant and usable. Implementing intelligent adaptation mechanisms can help optimize these interfaces, making them more responsive to user behavior and thus improving engagement levels. Abrahão et al. (2021) suggest that leveraging model-based adaptation techniques, such as machine learning, can dynamically adjust the interface based on user interactions, ensuring that the interface remains intuitive and user-friendly despite the flexibility it offers (Pedersen et al., 2018).

2.5. Challenges and Opportunities of Data Models and Interface Design

The relationship between data models and interface design is essential for understanding how systems manage data and facilitate user interaction. Data models provide the structural backbone, defining how information is organized, stored, and accessed. They establish rules for data relationships, constraints, and operations, ensuring system integrity, consistency, and reliability (Silberschatz et al., 1996; Parsons & Wand, 2000). However, while data models create the technical foundation, they must be translated into intuitive user interfaces to make these structures accessible and usable.

Interfaces serve as the bridge, transforming the abstract architecture of data models into user experiences. They enable users to perform tasks such as data entry, navigation, retrieval, and analysis—all informed by the underlying data model. When the interface aligns closely with the data model's architecture, users can engage with the system intuitively, minimizing confusion and enhancing the overall efficiency of data collection and management (Lukyanenko et al., 2018).

Designing interfaces for the IBDM presents unique challenges due to the model's flexible and dynamic nature. Unlike class-based models that rely on predefined categories, IBDMs capture unique data attributes, offering adaptability suited for dynamic environments like UGC or crowdsourced data (Lukyanenko et al., 2019; Parsons & Su, 2006). This flexibility, however, introduces complexity in interface design, as each instance is treated independently. Interfaces must manage variability and unpredictability in inputs, which can lead to inconsistencies if not designed carefully.

To address these challenges, research that focuses on developing adaptive interface designs capable of efficiently managing the dynamic characteristics of IBDMs is needed. Innovative tools and techniques are needed to enhance usability while minimizing cognitive load on users. Features such as personalized interfaces that adjust based on user behavior and input patterns can improve accessibility and efficiency (Parsons & Wand, 2000; Lukyanenko et al., 2018). Developing intelligent interfaces that can predict and adapt to user needs in real time is crucial for maximizing the advantages of IBDMs without overwhelming users.

Additionally, integrating IBDMs into existing frameworks requires balancing user-friendliness with system performance and data integrity. As data environments become more complex, ensuring interfaces handle sophisticated operations while remaining intuitive is essential. Achieving this balance will require ongoing research into adaptive design strategies that optimize user interactions while maintaining efficiency (Lukyanenko et al., 2019; Parsons & Su, 2010).

Chapter 3

3. Hypothesis Development

The development of the hypotheses is informed by the limitations identified in traditional Class-Based models and the potential advantages of Instance-Based approaches, as evidenced by previous studies.

3.1. Hypothesis 1 (H1)

IBID will capture more accurate and complete insights compared to traditional CBID, particularly by accommodating the variability and uniqueness of UGC.

This hypothesis is informed by the limitations of traditional Class-Based designs and the potential advantages of Instance-Based approaches, as evidenced by previous studies. In the context of UGC, users are data contributors, leading to different challenges in ensuring data quality. Traditional Class-Based models, which rely on predefined categories, can restrict the quality of UGC by constraining how users contribute information (Lukyanenko et al., 2014).

Research has shown that Class-Based modeling approaches can negatively impact IQ. Participants provide more accurate information when they can classify phenomena more broadly rather than being forced into predefined specific categories. Additionally, allowing users to provide free-form data improves overall accuracy compared to constrained choices (Lukyanenko et al., 2014). This supports the idea that IBID can capture more accurate and diverse user information by not relying on predefined categories.

Lukyanenko et al. (2018) further critique the dominant use of Class-Based models in conceptual modeling for ignoring the individuality of instances, leading to issues in accurately representing unique objects and their specific attributes. They propose Instance-Based modeling as a more effective approach for capturing the uniqueness of each instance. Theoretical and practical motivations support this approach, suggesting that IBID can better handle the uniqueness and variability of real-world data. Case studies show that IBID offers improved flexibility and support for unanticipated uses, facilitating novel discoveries (Lukyanenko et al., 2018).

Additionally, Lukyanenko et al. (2019) examined the impact of different data collection design choices on the quality of crowdsourced UGC. They found that Instance-Based data models result in higher accuracy and completeness by allowing contributors to describe phenomena using terms they are familiar with. This flexibility captures unexpected and novel aspects of data that Class-Based design might overlook.

By testing this hypothesis within a platform focused on student life experiences, I aim to extend the understanding of IBID beyond its previous applications in citizen science research. While prior studies (Lukyanenko et al., 2014, 2018, 2019) have demonstrated the benefits of Instance-Based models in improving data quality and completeness in citizen science contexts, this study is among the first to apply and evaluate these designs in a broader UGC environment.

3.2. Hypothesis 2 (H2)

IBID enables more authentic user expression than CBID.

This hypothesis is informed by several key findings from past research that highlight the limitations of traditional Class-Based models and the potential advantages of Instance-Based approaches. Class-Based models often fail to represent the individuality of instances, leading to issues in accurately capturing unique objects and their specific attributes. This limitation is critical in the context of dynamic and varied UGC. Forcing contributors to conform to predetermined categories in Class-Based designs introduces bias, as users may select categories that do not accurately reflect their observations. This bias occurs because the available predefined options may not fully align with the contributor's understanding of the data, which results in incorrect classification and information loss (Lukyanenko, 2014). For example, empirical studies show that users provide more accurate responses in free-form data entry, achieving higher classification accuracy compared to schema-mediated tasks where predefined categories are imposed. The presence of predefined options can mislead users into making choices they might not otherwise make, ultimately reducing the quality of the reported data (Lukyanenko, 2014).

IBID, on the other hand, emphasizes the primary role of instances, allowing for unique and unconstrained data entry, which can better capture the authenticity of user contributions (Lukyanenko et al., 2018). This approach eliminates the constraints imposed by structured interface designs, enabling users to report their observations more freely and accurately. By focusing on individual instances rather than predefined categories, IBID allows for a more accurate and complete representation of UGC, capturing unexpected and novel aspects of data that Class-Based models might overlook (Lukyanenko, 2014). This flexibility enables more authentic user expression, supporting the hypothesis that IBID can improve data quality by removing biases caused by predefined choices (Lukyanenko et al., 2019).

Given these findings, this study hypothesizes that Class-Based interfaces bias user reporting by imposing predefined categories that constrain user expression.

3.3. Hypotheses and Research Focus

In summary, the main two hypotheses are:

H1: IBID will capture more accurate and complete insights compared to traditional CBID, particularly by accommodating the variability and uniqueness of UGC.

H2: IBID enables more authentic user expression by removing the biased user reporting constraints.

This study aims to test these hypotheses by comparing the performance of Class-Based and Instance-Based interface designs in a platform focused on student life experiences. By doing so, I seek to extend the understanding of Instance-Based modeling and its practical benefits in enhancing data quality and user engagement in the context of UGC. Lukyanenko et al. (2014, 2018, 2019) researched the impact of conceptual modeling on UGC's IQ. They argued that traditional Class-Based models, which use predefined categories, restrict UGC quality by limiting how users contribute information. Their studies showed that participants provide more accurate data when using general classifications or free-form interface designs instead of predefined categories. This supports the idea that IBID can capture more accurate and diverse information. They suggested reengineering conceptual modeling to focus on Instance-Based representations, which better capture the uniqueness of each instance and improve UGC accuracy and completeness. This approach allows for unique, unconstrained data entry, leading to higher-quality crowdsourced UGC.

Moreover, previous research has primarily focused on citizen science to test their hypotheses. Unlike citizen science, where data is externally verifiable and unpredictability stems from environmental factors, student life reporting deals with subjective, personal experiences that lack predefined categories. Additionally, while citizen science participants engage out of scientific interest, student respondents may participate due to institutional requirements or personal concerns, affecting engagement and response depth. This study highlights IBID's adaptability in capturing authentic, diverse insights in dynamic social contexts.

Chapter 4

4. Experiment Design

The experiment was designed using a randomized controlled trial (RCT) model. A randomized controlled trial is a prospective, comparative, quantitative study performed under controlled conditions with random allocation of interventions to comparison groups testing the effectiveness and/or safety of one or more interventions (Bhide et al., 2018). To facilitate the execution of the experiment, I developed two distinct interface designs for comparative analysis. The initial interface adhered to the conventional principles of CBID, serving as the control group throughout the course of the experiment. This choice was grounded in the well-established tradition and widespread utilization of the CBID within the field.

Conversely, the second design was rooted in unstructured interface design, and it was deliberately selected as the experimental treatment. The primary objective was to investigate the potential effects and repercussions that this relatively novel and less widely adopted design could have on the data collection processes, particularly regarding acquiring unforeseen or unanticipated data. This approach enabled us to compare traditional and novel interface designs comprehensively under controlled experimental conditions. The survey content is an accumulated effort of previous research findings and the author's judgment. Further explanation of this process can be found in more detail in section 4.2.

To diversify the demographic focus from previous research, which has primarily centered on citizen science, this study specifically targets students at Memorial University of Newfoundland (MUN). To achieve a comprehensive demographic representation, the participant pool includes students from various academic programs and levels across the campus. Recognizing the significance of the target population and sample selection for the experiment, MUN students were chosen for their diversity in academic backgrounds and experiences, providing a well-rounded basis for the study's findings.

To facilitate participation, I contacted potential participants via an anonymous email invitation sent to their MUN email addresses. This invitation contained a link to the survey, allowing students to anonymously submit their responses and contribute to the research. To encourage participation, a random draw was conducted, in which four respondents were awarded $50 gift cards. This incentive aimed to increase engagement and response rates while maintaining voluntary participation. Upon completing the study, the winning participants received their gift cards.

4.1. Class-Based Interface Design

In designing the CBID, I explored various resources and references to ensure the survey was thorough and relevant. Although finding extensive resources was challenging due to the niche nature of the topic and the confidential nature of many existing surveys, I drew inspiration from available sources. These included academic surveys conducted by other universities, such as the Independent Student Analysis Report from the University of Toronto, and other relevant documents like Students' perception of online learning during the COVID-19 pandemic. Additionally, I utilized Memorial University's Institutional Survey Oversight Committee (ISOC) Question-Wording Templates and Considerations for Question Design to ensure the questions were appropriately formulated.

To establish the initial categories for the survey, I used these references to identify three categories that appeared most frequently in the literature. However, after conducting a pilot interview, I refined these categories to ensure their coherence with the questions they contained. The result included three main categories: academic experience; safety, inclusivity, and well-being; and facilities.

To provide a clear overview of how different questions were inspired by and mapped to these sources, I have included Table A.1 in Appendix A, Mapping the Survey Questions to Reference Sources. This table outlines the mapping of each question to the specific references that informed their creation. The final version of the questions was further edited and customized to suit the specific context of the survey, with the resources serving as the building blocks and initial guidance for the survey's content.

After drafting the initial survey, pilot testing was conducted through interviews with six Memorial University students. This pilot phase was essential in refining the survey, specifically to enhance the validity of the content in the class-based survey and serve as a building block for the predefined classes. The interviews involved a diverse group of participants from various programs, ensuring that the feedback gathered was representative of the broader student population. During this process, several questions were edited or adjusted to better align with the participants' experiences and the context of Memorial University. For instance, questions drawn from the University of Toronto's survey were initially numerous, and to avoid overwhelming participants and potentially discouraging completion, I decided to reduce the number of questions. Based on the interview feedback, I removed some questions that were deemed irrelevant and tweaked others to ensure they were more relatable and accessible to Memorial University students.

Moreover, the interviews revealed nuances not initially captured in the survey design. As a result, I made specific edits to questions and added new ones to address emerging themes, ensuring that the class-based survey categories accurately reflected student perspectives. These adjustments helped to ensure that the survey was not only comprehensive but also concise and user-friendly, balancing the need for detailed feedback with participants' willingness to complete the survey. Examples of specific questions that were added, edited, or removed based on interview feedback can be found in Appendix B.

The CBID was designed to present participants with three primary categories: academic experience, safety, inclusivity, and well-being, and facilities. Depending on their selection, participants were directed to the corresponding set of questions structured as statements, which they were asked to rate using a Likert scale. Additionally, a text box was provided at the end of each form, allowing respondents to share any further comments. The structure of the CBID, as well as the Entity-Relationship Diagram (ERD), can be found in Appendix C.

4.2. Instance-Based Interface Design

Participants assigned to the IBID condition encountered a more flexible and unstructured data input interface. Unlike the structured approach of the CBID, the IBID allowed for greater flexibility in how data could be entered, reflecting the less rigid nature of Instance-Based data management systems.

Participants were instructed to input data as they would in a free-form document or an open-ended form. The interface allowed for unstructured Text Entry. Participants could input data in narrative form, using full sentences or bullet points, without being constrained by predefined fields. To ensure participants understood how to use the flexible interface, detailed instructions and examples were provided at the beginning of the survey.

4.3. Survey Flow

The data collection process commenced with participants entering the survey through anonymous links provided to them. Upon consenting to the privacy statement, they were presented with the demographic questions. After completing these questions, participants were randomly assigned to either the IBID or CBID. This random assignment was ensured through the Qualtrics process flow, detailed in Figure 4.1.

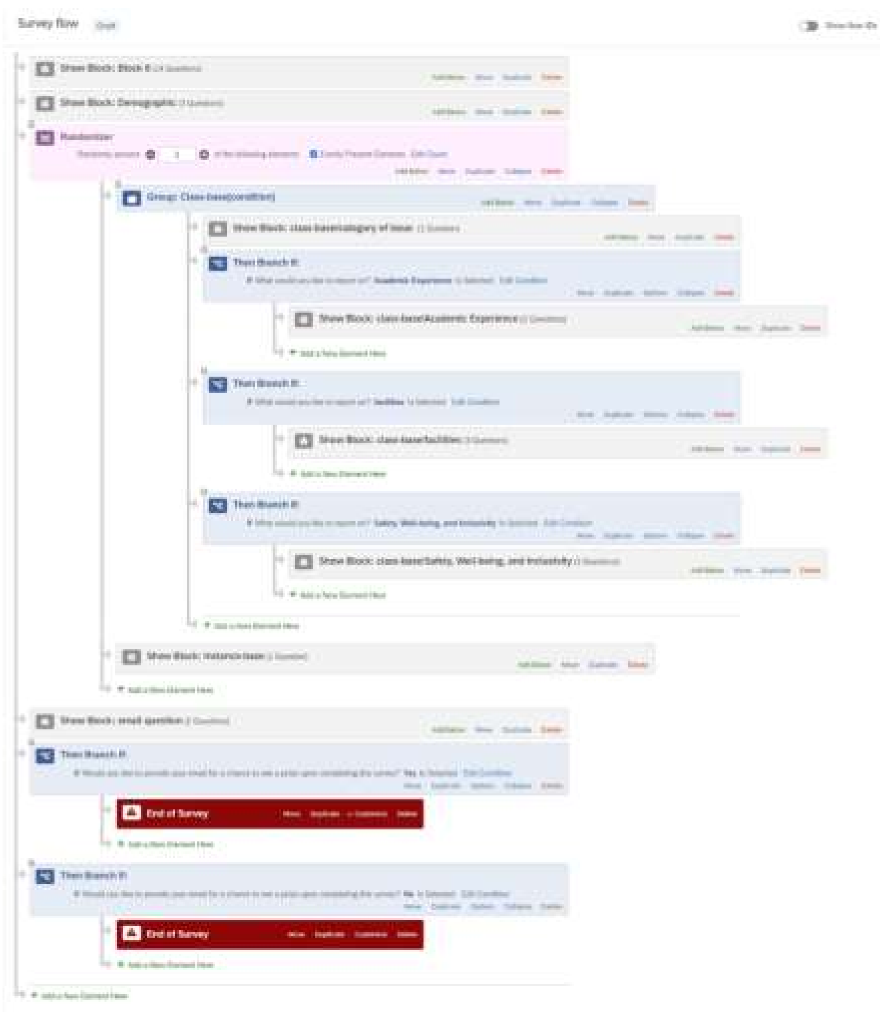

Figure 4.1 ERD Survey Flow

Participants assigned to the CBID were presented with three categories of issues to choose from: academic experience; safety, inclusivity, and well-being; and facilities. After selecting a category, they completed the corresponding form. Their responses were recorded only after they submitted the survey, ensuring that participants could leave the survey at any point if they chose to. Similarly, participants assigned to the IBID filled out the IBID form, and their responses were recorded upon submission.

Qualtrics was utilized to collect data. The platform's capabilities allowed us to collect not only participants' responses but also metadata, such as the duration each participant took to complete their data entry, providing insights into the ease of use and effort engaging with the interface. Detailed instructions and guidance were provided to participants to ensure they completed the tasks appropriately, emulating real-world data entry scenarios relevant to each interface design. The structure and content of the survey can be found in the Appendix D.

Chapter 5

5. Results

This chapter presents the findings and analysis of the collected data.

5.1. Data Collection

From April 22nd to June 10th, 2024, I collected responses from students using the Qualtrics platform. During this period, 83 responses were submitted. Participants were presented with one of the two interface designs: Class-Based or Instance-Based. Each design was shown 59 times, a total of 118 presentations.

Out of the 118 presentations, I removed the submitted responses that did not contain any data in the Class-Based or Instance-Based design from the recorded responses. Thereafter, I had 30 responses for the Instance-Based design and 48 responses for the CBID.

5.2. Distribution of Academic Degrees

The distribution of academic degrees among the respondents indicates a diverse range of educational qualifications. The majority of the respondents hold either a bachelor's or master's degree, with 32 and 28 individuals, respectively. The distribution of academic degrees is visually represented in *Figure 5.1*. Moreover, refer to Table E.1 in Appendix E for detailed numerical data.

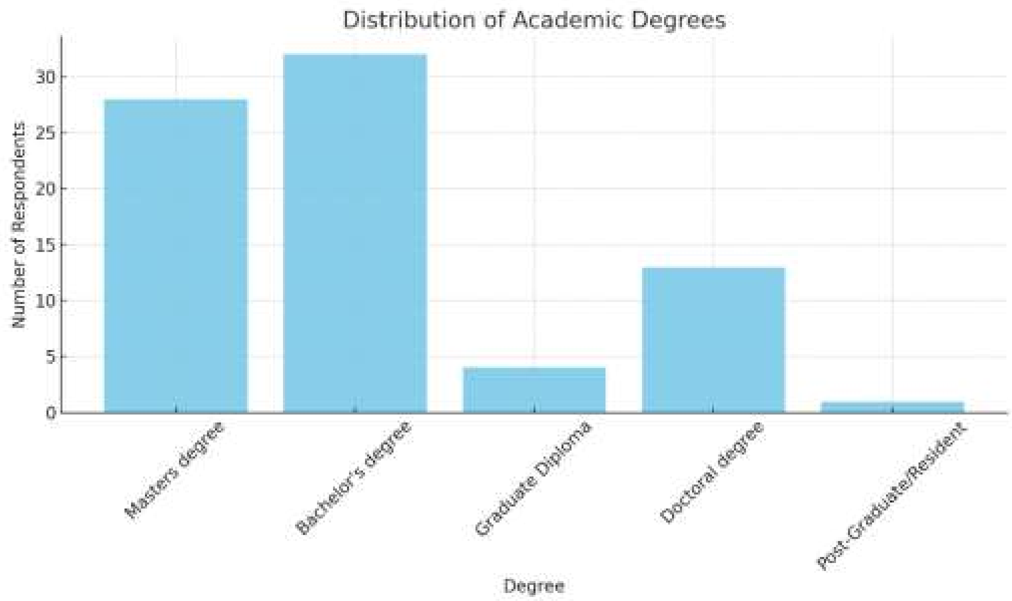

Figure 5.1 distribution of academic degrees

5.3. Distribution of Academic Programs

The distribution of academic programs among the respondents reveals a broad spectrum of disciplines. The largest group is from Business Administration, with 39 respondents. This is followed by Science with 11 respondents, and Engineering and Applied Science with 10 respondents. Other programs such as Humanities and Social Sciences (5), Education (3), and various other disciplines show a wide array of academic interests. The detailed numerical data for the distribution of program years can be found in

Table E.2 in the Appendix E. Additionally, this distribution is visually illustrated in *Figure 5.2*.

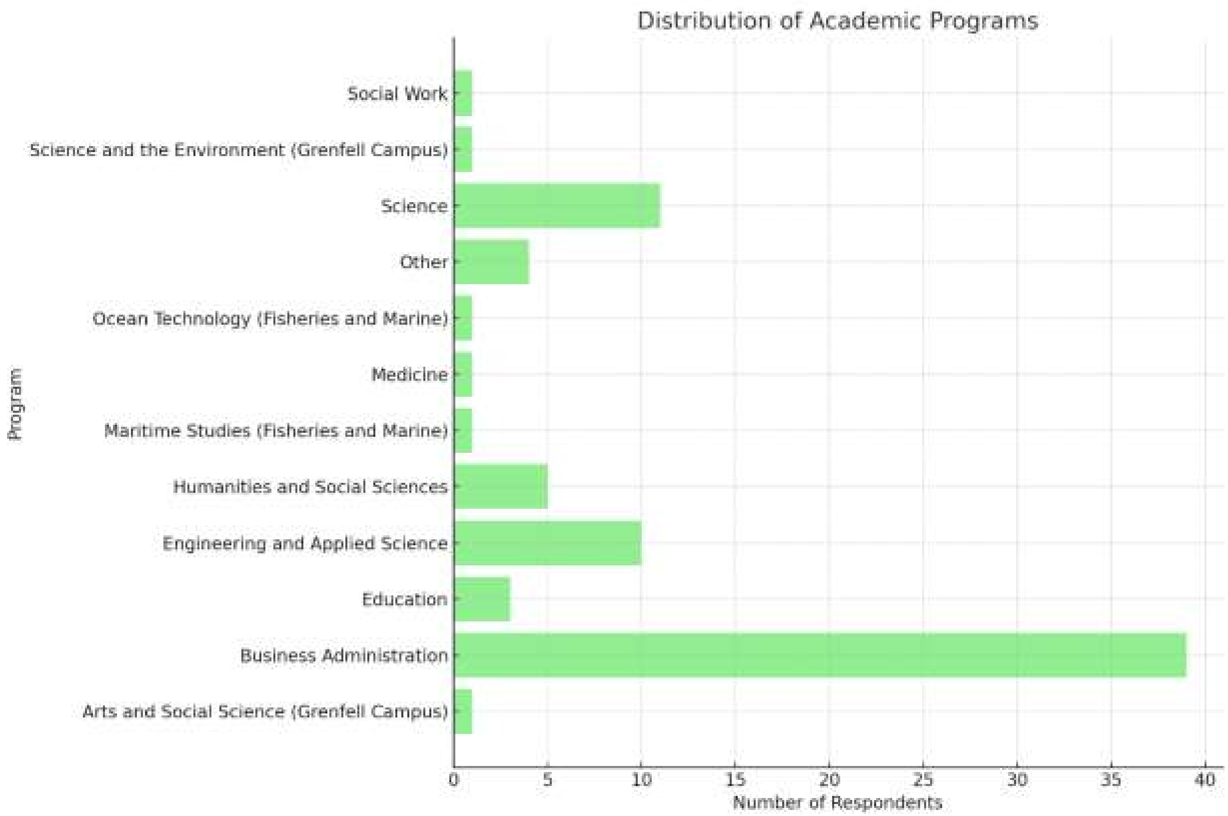

Figure 5.2 Distribution of program years

5.4. Distribution of Program Years

The distribution of respondents by their program year indicates a varied representation across different stages of their academic journey. For detailed numerical data, refer to

Table E.3 in the Appendix E. This distribution is visually illustrated in *Figure 5.3*.

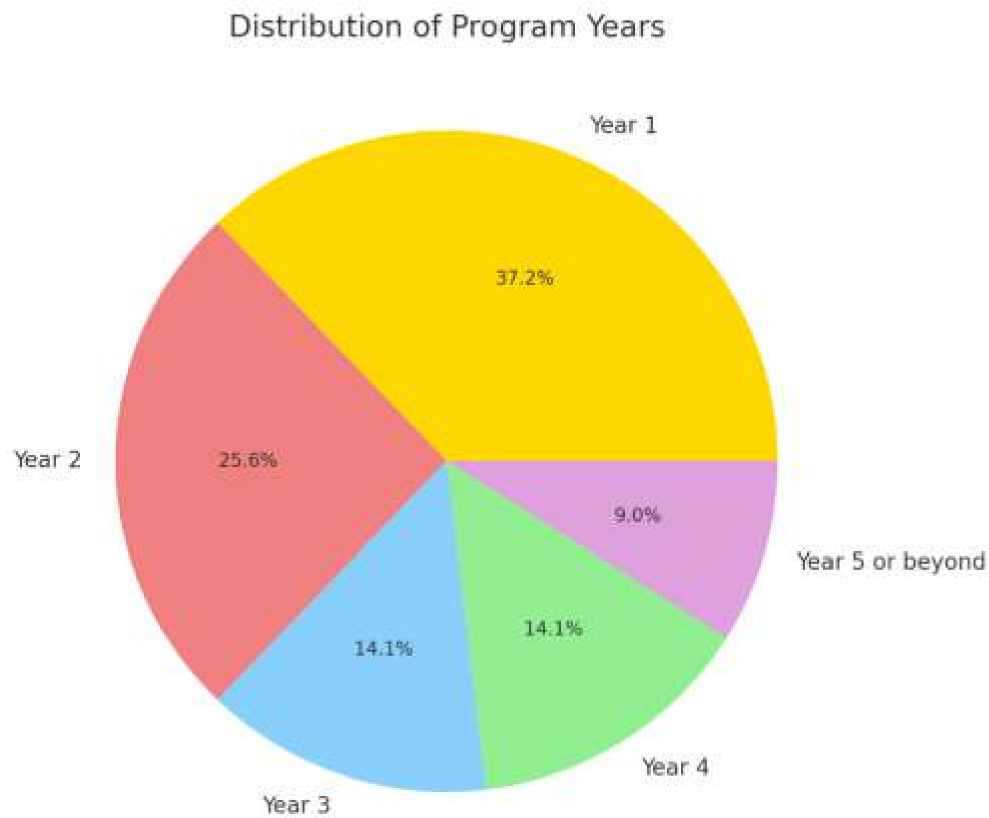

Figure 5.3 Distribution of Program Years

5.5. CBID Responses

I collected 48 responses using CBID to gain insights into the effectiveness and utility of the two interface designs. Initially, demographic information was gathered from the respondents. Subsequently, they were asked to categorize their reports into one of three categories: Safety, Well-being, and Inclusivity; Facilities; or Academic Experience.

The overwhelming majority of responses focused on "Academic Experience," with 85.4% of respondents selecting this category, while only a few responses addressed Facilities (6.3%) or Safety, Well-being, and Inclusivity (8.3%). The distribution is illustrated in Figure 5.4.

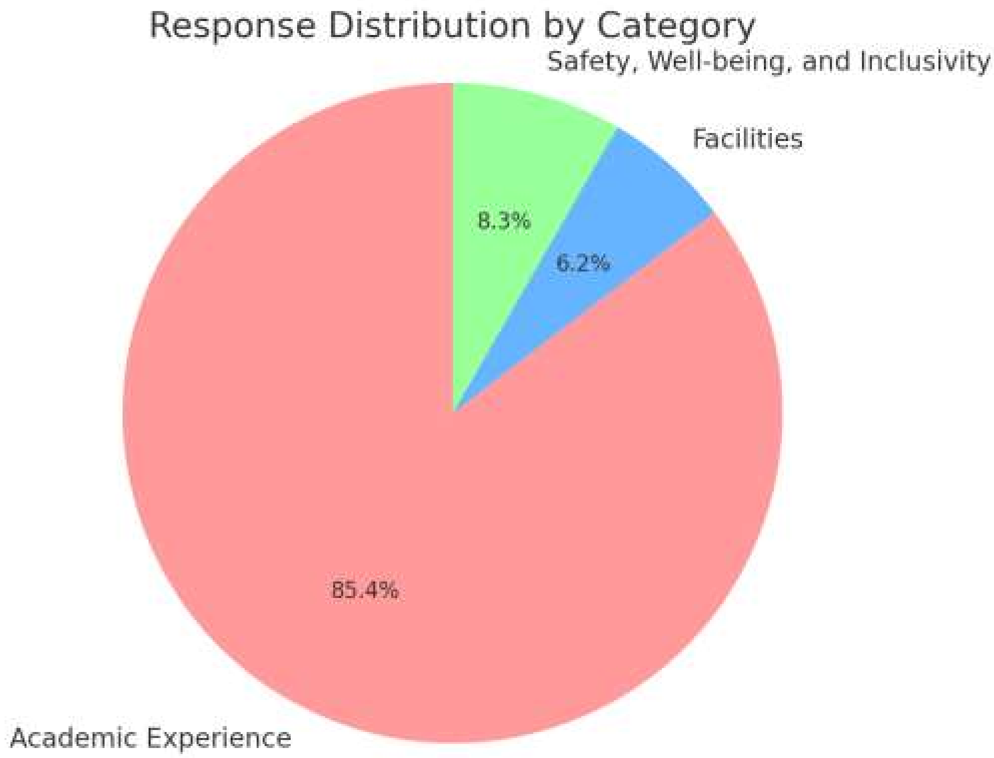

Figure 5.4 Response Distribution by Category

After selecting a category, respondents were presented with a series of statements and asked to indicate their level of satisfaction on a Likert scale. Each category form concludes with a text box, allowing users to provide further details if they choose to do so.

Only eight respondents utilized the text boxes to provide additional comments. Upon closer inspection, one of these responses did not contain any valuable information, leaving us with seven substantive responses in the text boxes. Additionally, the eight responses in the text boxes had an average word count of 26.5 words.

Lastly, the average duration for completing the survey was 204.5 seconds.

5.6. Instance-Based Design Responses

Unstructured data collected through the IBID condition required a different approach to analysis compared to the structured data from the CBID condition. Methods for analyzing this data are thoroughly explained below.

5.6.1. Quantitative analysis

To begin the quantitative analysis, I examined two key metrics: the time taken to complete the IBID survey and the word count of the responses.

5.6.1.1. Response Time Analysis

The average time for respondents to complete the IBID survey was 430.03 seconds. This data suggests that the Instance-Based design requires more time from participants due to the open-ended nature of the responses. *Figure 5.5* shows the distribution of participant durations grouped into 100-second intervals.

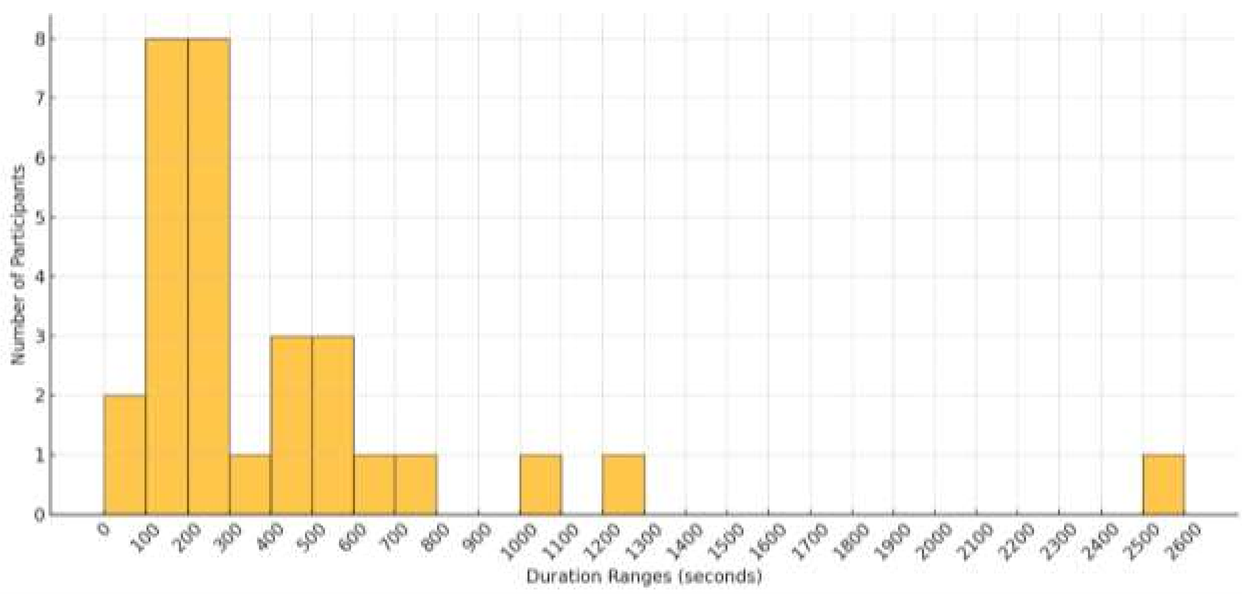

Figure 5.5 Duration of Responses versus Submissions

5.6.1.2. Word Count Analysis

Next, I analyzed the word count of each response in the IBID text box. Appendix F contains the code used for this function, which was determined using the Visual Basic for Applications (VBA) feature in Microsoft Excel). On average, users submitted 55.90 words in the instance box. *Figure 5.6* illustrates the word count of each response, the distribution of word counts among participants, grouped into intervals of 10 words, with the y-axis representing the number of participants in each range.

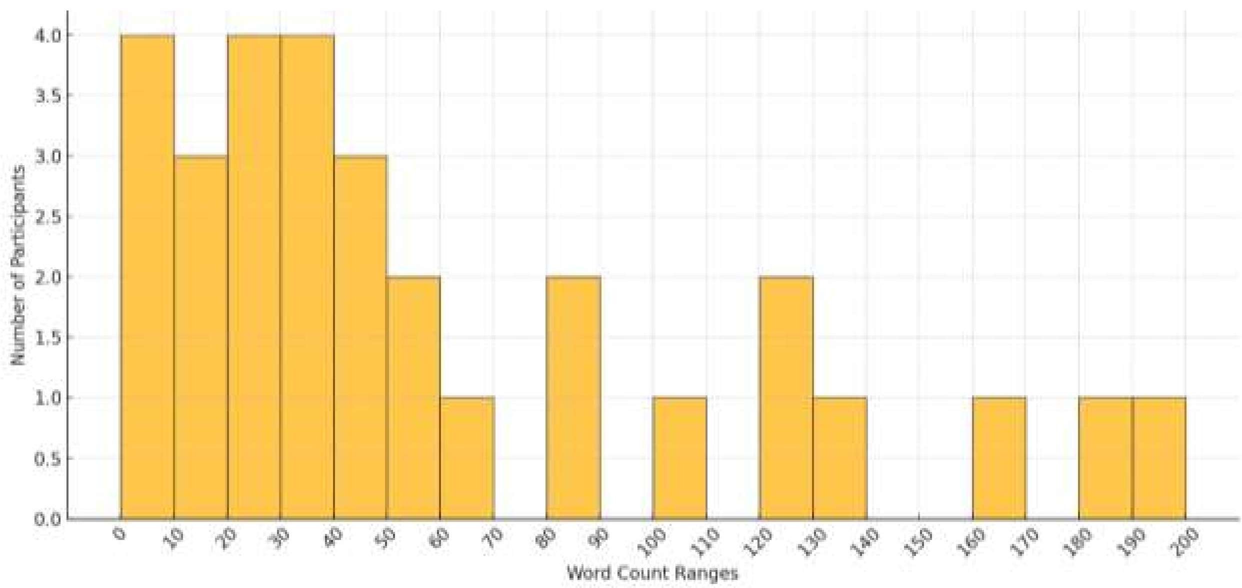

Figure 5.6 Word Count of Responses versus Submissions

5.6.2. Qualitative Analysis

The qualitative analysis of the IBID responses aimed to uncover patterns, sentiments, and insights that the CBID may not have captured. This analysis was conducted using two approaches: thematic analysis and sentiment analysis.

5.6.2.1. Thematic Analysis

I used thematic analysis to analyze the responses and further investigate the use of the two different interface designs. Braun and Clarke (2006) define thematic analysis as a method for identifying, analyzing, and reporting patterns or themes within data. This approach enabled me to categorize the IBID responses into the three initial groups. I selected the initial three categories from the CBID responses as the framework for the IBID responses to gain better insights into the comparative investigation between the two interface designs. The analysis process included theming the data into categories, reflecting and synthesizing the data, and condensing data to assist in reporting and interpreting the data and the findings (Brenner, 1985). The analysis was conducted using Microsoft Excel.

The first step involved assessing the relevance of each response to student life experiences. Out of the 30 responses, 28 were found to be relevant, while two were not. To determine relevance, I manually scanned the text of each response, looking for specific references to aspects of student life at MUN. The two irrelevant responses provided only general life experiences in Canada and did not elaborate on any specific aspect of student life at MUN.

Next, the 28 relevant responses were categorized into four groups: Safety, Wellbeing, and Inclusivity; Academic Experience; Facilities; and Others. This categorization was done by reading and interpreting the underlying points conveyed in the texts submitted via the Instance-Based design text box. I successfully categorized 26 responses into the initial three groups, while four responses did not fit any of these categories and were placed into a fourth category labeled "Others". For example, one respondent suggested that newly admitted students should receive guidance on study permit applications and emphasized the need for more affordable housing options, especially for international students. This response did not directly align with the existing categories as it touched on broader institutional support needs. As a result, it was categorized under "Others."

Additionally, four responses were categorized into two groups as they provided insights into both categories. Specifically, I matched 14 responses with the "Safety, Wellbeing, and Inclusivity" category, 10 in the "Academic Experience" category, and 4 in the "Facilities" category. *Figure 5.7* shows a visual presentation of the analysis process.

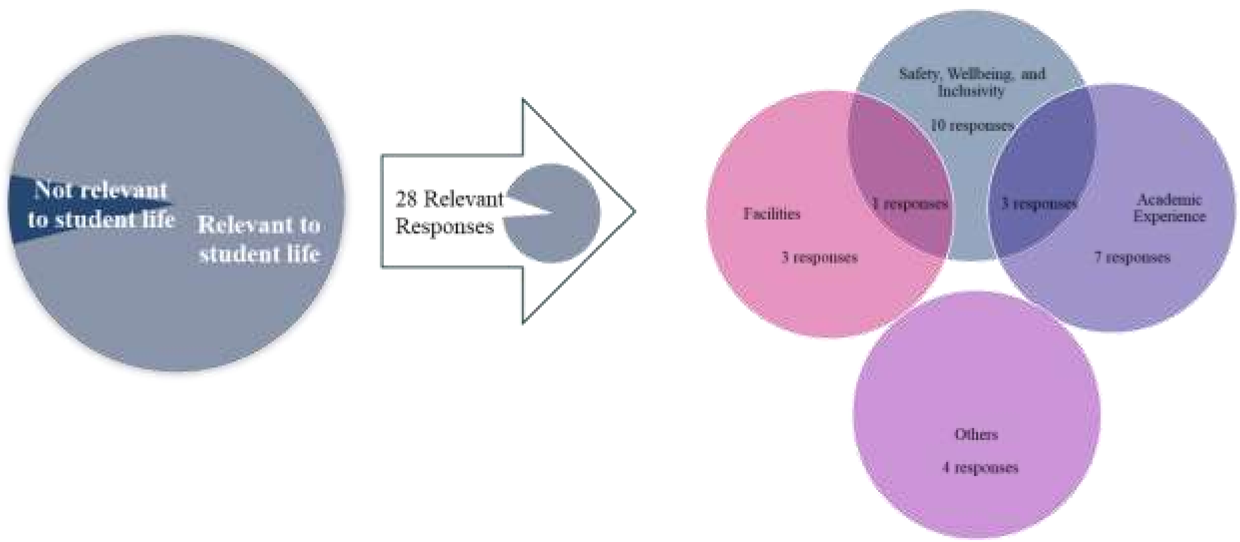

Figure 5.7 Categorization of Relevant Responses

Moreover, I analyzed whether the insights provided in the relevant responses were included in the CBID. In other words, I aimed to determine if the information gathered from the responses was also covered in the CBID. To do this, I manually coded each submitted text. I read each response to identify whether the questions in the CBID could have captured the insight and value the submitted IBID response provided.

For example, one IBID response indicated that the student was always able to seek support from university staff. I checked if this insight could be derived from any of the questions in the CBID. The information could be gathered through the questions: “Rate the visibility and responsiveness of security personnel in bias-related incidents" and "Rate how well the campus communicates available safety, well-being, and inclusivity resources." Since the information could be gathered through these questions, I counted this response as one of the 12 records accounted for in the CBID. The remaining 16 responses provided insights not covered by the CBID, indicating that, despite the efforts to encompass possible insights from students and suggestions from pilot interviews, more than half of the information provided by the submitted reports was not accounted for in the CBID design.

Finally, I examined the reports that provided information included in the CBID to determine whether they also offered insights beyond the questions asked in the CBID. From the 12 submissions that were accounted for in the survey, I took an additional step to investigate whether these comments provided any further insights that could not be captured through the CBID.

Out of the 12 responses, 8 provided additional information that went beyond the scope of the CBID, while 4 did not. For example, one response captured within the CBID discussed the availability of instructors and the appropriateness of course outlines, as well as insights already included in the CBID. However, the respondent went on to describe the lack of clear communication regarding whether online courses would require in-person exams or necessitate logging in from St. John's for several days. This level of detail could not have been gathered through the existing CBID questions, and thus, it was counted as providing additional insights. Through this analysis, eight responses were identified as contributing further insights beyond the scope of the CBID. This process is depicted in Figure 5.8.

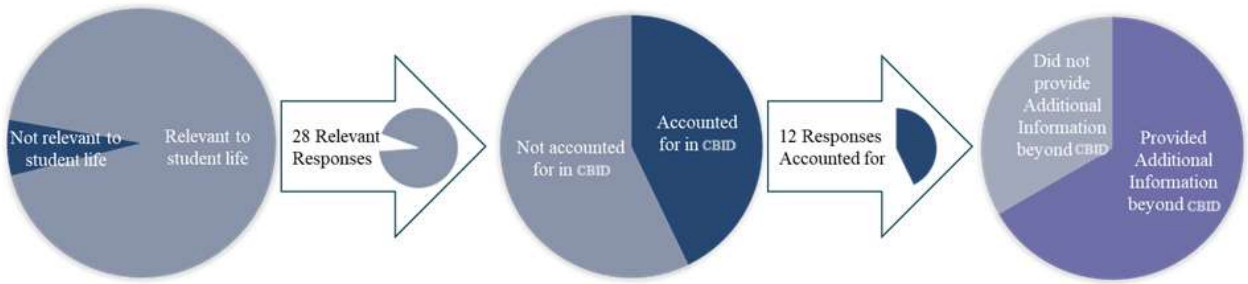

Figure 5.8 Analysis Process for Comparing IBID versus CBID in Capturing unanticipated data

5.6.2.2. Sentiment Analysis

Sentiment analysis is a qualitative research method used to assess the emotional tone or sentiment expressed in responses, enabling researchers to categorize feedback as positive, negative, or neutral (Pang & Lee, 2008). In this study, sentiment analysis was conducted to evaluate the emotional undertones of the feedback collected through the IBID. The primary aim was to categorize the responses and compare the effectiveness of the IBID with the CBID in capturing a broader range of sentiments.

The IBID was designed to collect more open-ended and diverse input from students, allowing for a broader range of sentiments compared to the structured CBID, which constrained responses to positive or negative categories. The sentiment analysis aimed to demonstrate how the IBID broadened the scope of feedback by capturing not only satisfaction and dissatisfaction but also constructive improvement suggestions and, occasionally, non-informative responses.

For the analysis, I categorized each response into the following groups:

- Positive Statement: Responses expressing satisfaction or highlighting positive aspects of student life at MUN.
- Negative Statement: Responses reflecting dissatisfaction or highlighting negative aspects of student life at MUN.

- No Value-Added Statement: Responses that fail to provide relevant or actionable information.
- Improvement Note: Responses offering constructive feedback or suggestions for improving student life.

The CBID constrained student feedback to primarily positive and negative responses, limiting the range of insights that could be gathered. In contrast, the IBID allowed for more nuanced input, acting as a double-edged sword. On the positive side, the IBID captured Improvement Notes, which were highly valuable for providing actionable insights and understanding what students believe should change. On the downside, it also captured No Value-Added Statements, which provided little relevant information.

Overall, the IBID's flexibility enabled a more detailed and constructive form of feedback, which provided deeper insights into the diverse nature of student experiences. While some non-informative feedback was collected, the ability to gather valuable suggestions outweighed this limitation. After categorizing the 30 IBID responses, I found the following distribution: 11 responses were positive statements, 11 were negative statements, 3 were no value-added statements, and 5 were improvement notes. *Figure 5.9* depicts this distribution visually.

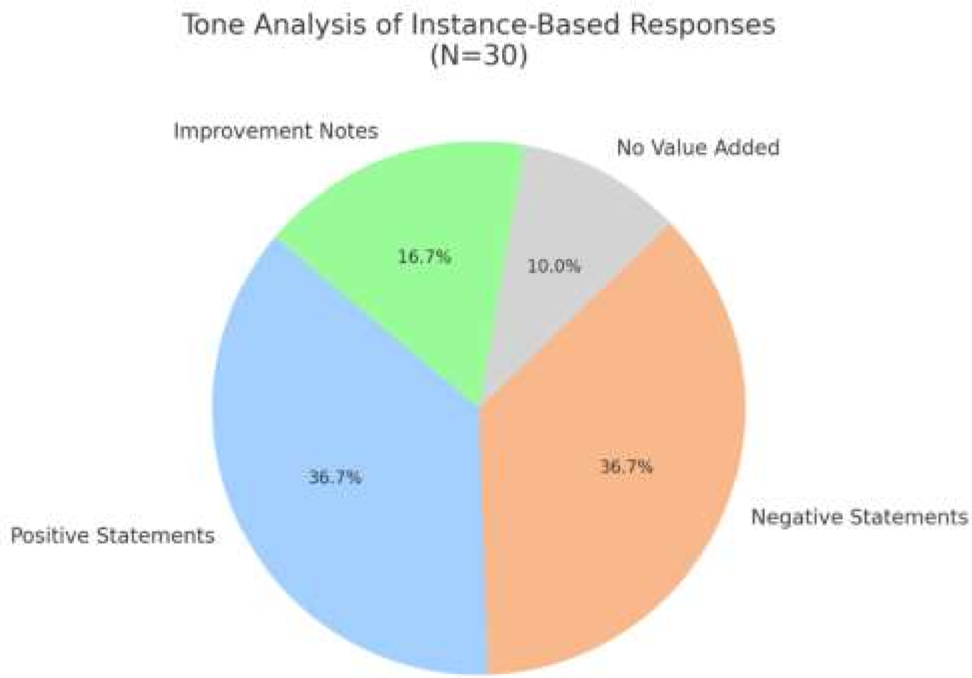

Figure 5.9 Distribution of Sentiment in IBID Responses

5.6.3. Artificial Intelligence (AI) Thematic Analysis

The thematic analysis was conducted using ChatGPT-4.0, an AI language model developed by OpenAI. This approach helped to systematically analyze the survey responses, providing valuable insights into the experiences and concerns of students at MUN.

The following prompts were used in the AI thematic analysis:

- "Analyze survey responses to identify recurring themes."
- "Categorize the feedback based on common areas of concern and positive aspects."
- "List and explain the identified themes, focusing on specific issues and their impact."

This thematic analysis categorizes the survey responses into distinct themes based on the issues mentioned. The analysis identifies common areas of concern and positive feedback, helping to pinpoint specific issues and areas for improvement. Below is a summary of the analysis, including the number of responses in each category and a brief explanation of each theme. Moreover, Figure 5.10 depicts this distribution visually. The finding of this analysis is further explored in section 6.6.2.

Categories and Number of Responses:

- Academic Experience (5 Responses)
- Student Life (3 Responses)
- Housing and Residence (4 Responses)
- Administrative Issues (3 Responses)
- Support Services (3 Responses)
- Financial Concerns (3 Responses)
- COVID-19 Impact (2 Responses)

Category Explanations:

Academic Experience (5 Responses)

This category includes responses related to the overall educational experience at MUN, including course availability, learning opportunities, and support from academic staff. Respondents mentioned both positive experiences and areas needing improvement in their academic journey.

Student Life (3 Responses)

This category covers aspects of student life outside of the classroom, including social activities, involvement in student societies, and the general atmosphere on campus. The impact of COVID-19 on student engagement and social events is a significant concern highlighted by respondents.

Housing and Residence (4 Responses)

Responses in this category address issues related to on-campus housing and off-campus living conditions. Students expressed dissatisfaction with the quality of on-campus residences, high rent prices, and the availability of housing. Parking issues for students living on campus are also noted.

Administrative Issues (3 Responses)

This category encompasses problems related to university administration, such as a lack of transparency and communication in interdisciplinary graduate programs and the need for better program planning and development. Respondents reported missing important deadlines and receiving inadequate support from administrative staff.

Support Services (3 Responses)

Responses in this category highlight the need for better support services for students, including assistance with co-op placements, mental health support, and general academic support. Students appreciated the supportive community but noted areas where additional help is needed.

Financial Concerns (3 Responses)

This category focuses on financial issues faced by students, including the high cost of living, expensive food options on campus, and the perception of the university operating more like a business than an educational institution. The high costs associated with studying at MUN are a significant concern.

COVID-19 Impact (2 Responses)

This category includes responses specifically mentioning the impact of the COVID-19 pandemic on their student experience. Issues include the shift to remote learning, the effect on student engagement, and the challenges of adjusting to new teaching methods and balancing family/work/school responsibilities.

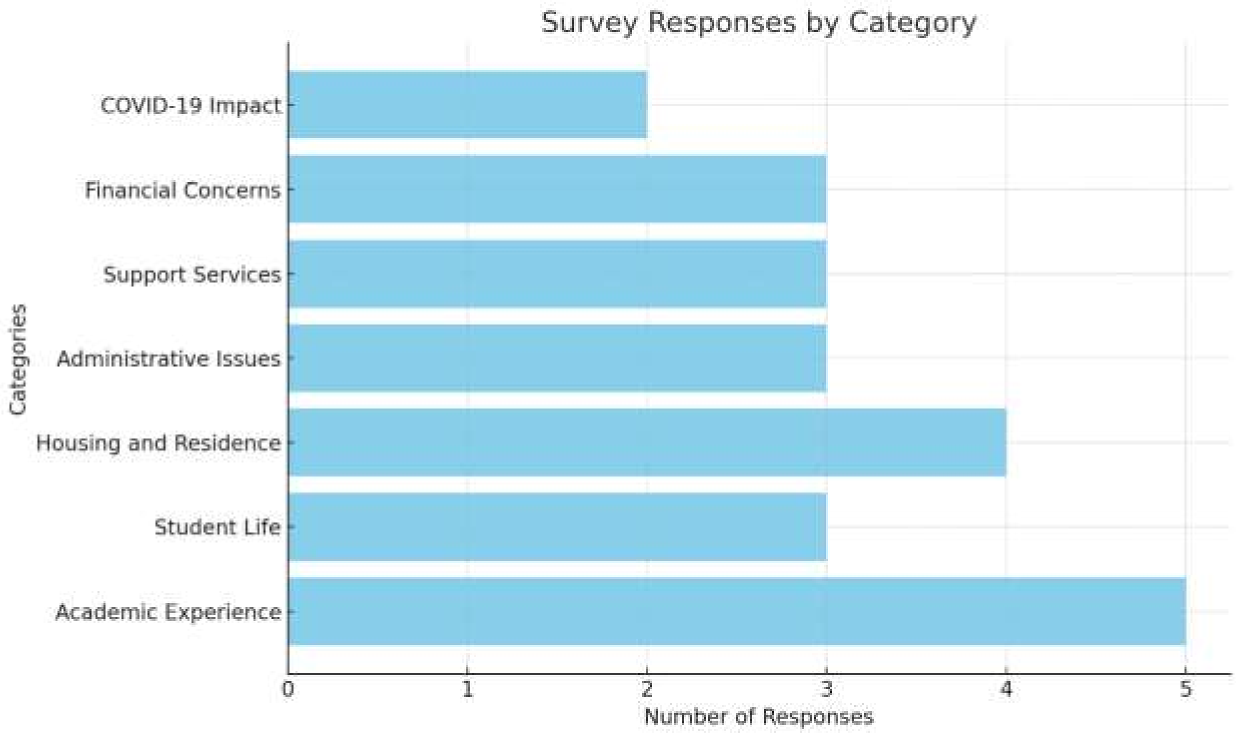

Figure 5.10 Survey Responses by Category through AI

Chapter 6

6. Discussion

This section discusses the findings of the analysis and compares the effectiveness of the Instance-Based design versus the Class-Based design.

6.1. Survey Abandonment Rates

Out of the 118 presentations, each design was shown to 59 participants, ensuring equal opportunity for response. However, a significant difference in completion rates was observed: 48 responses were recorded for the Class-Based Information Design (CBID) and 30 responses for the Instance-Based Information Design (IBID). This indicates that 29 respondents abandoned the survey when presented with the IBID, while only 11 abandoned the survey when shown the CBID. A chi-square test for independence revealed a statistically significant difference in abandonment rates between the two designs ($\chi^2 = 10.93$, $p < 0.001$). These results demonstrate that respondents were significantly more likely to abandon the survey when exposed to the IBID compared to the CBID. The disparity in response and abandonment rates between the two designs is noteworthy. The higher abandonment rate for the IBID suggests potential usability or comprehension issues compared to the CBID. This discrepancy may be attributed to the nature of the IBID, which requires respondents to provide free-form text inputs. Such an open-ended approach might be perceived as more time-consuming and demanding, leading to a higher likelihood of survey abandonment.

In contrast, the CBID structures responses within specific categories and employs Likert scale ratings. This structured format appears to be more user-friendly, as it likely reduces the cognitive load on respondents and provides a clearer path to completion. The ease and speed of completing the survey with the CBID may contribute to its higher response rate.

This analysis clearly indicates that while the IBID offers flexibility and a wide range of uses, it may not be the optimal choice in scenarios where user motivation to finish the survey is crucial. The structured approach of the CBID seems to encourage higher completion rates, suggesting it might be a more effective and user-friendly option in such contexts. This finding underscores the importance of considering user experience and survey design in research methodologies to ensure high response rates and reliable data collection.

6.2. Survey Duration Comparison

In this section, I compare the survey completion time between CBID and IBID. Respondents in the CBID took an average of 204.5 seconds. In contrast, respondents using the IBID spent an average of 430.03.

This data reveals that respondents using the IBID invested more than double the time compared to those using the CBID. This increased duration can be attributed to the nature of the IBID, which allows respondents to express their opinions freely in a text box rather than responding to predefined questions. This free-form format requires respondents to think more deeply about their responses, leading to more time spent on the survey. Moreover, the extended duration for the IBID suggests that respondents engaged more deeply with the survey, taking time to reflect and provide detailed responses. Consequently, it is likely that the IBID elicits more detailed and nuanced answers compared to the CBID. This format is advantageous for obtaining in-depth qualitative data, especially when time constraints are not a critical factor.

However, the increased time requirement also highlights a potential drawback. For surveys where time is limited or where respondents may not be willing to invest a significant amount of time, the IBID may not be the optimal choice. In such cases, the CBID, with its quicker response time, would be more suitable. Its structured format facilitates quicker completion, making it suitable for large-scale surveys where brevity and ease of response are prioritized.

6.3. Word Count Analysis

A second key point of comparison is the average word count of the text box responses in each interface design. In the IBID, the average word count was 55.9 words, whereas in the CBID, the average word count was 26.5 words. This indicates that respondents provided more than double the amount of text in the IBID compared to the CBID.

This difference in word count suggests that users felt more compelled to provide detailed responses in the IBID. The open-ended nature of the IBID may have encouraged respondents to elaborate more on their experiences and thoughts. In contrast, the structured format of the CBID may have led respondents to feel that additional textual input was unnecessary, or they may have been less inclined to provide lengthy responses due to the predefined response options.

Furthermore, it is noteworthy that only 8 out of 48 respondents in the CBID chose to provide additional text in the provided text boxes. This low rate of additional comments suggests that the structured questions in the CBID were perceived as sufficient for capturing the necessary information, thereby reducing the respondents' motivation to add their unique experiences or further details. Another possible reason for this outcome could be the influence of anchoring.

Anchoring is a cognitive bias where individuals rely heavily on the initial information presented to them, using it as a reference point (the "anchor") in their decision-making process (Lieder et al., 2017). In this context, the structured questions and response options in the CBID serve as an anchor, signaling to respondents that detailed text comments may not be necessary. This anchoring effect can lead to a reduced inclination to provide additional information as respondents align their responses with the structured format, assuming that the initial questions sufficiently cover the required insights.

This comparative analysis highlights that, while the CBID may facilitate higher response rates and quicker survey completion, the IBID encourages more detailed and comprehensive textual feedback. Researchers must balance these considerations when choosing between these designs, depending on the specific objectives of their data collection efforts. For instance, if detailed qualitative data is essential, the IBID might be more appropriate. Conversely, if higher response rates and structured data are the priority, the CBID may be more suitable.

6.4. Insights from Response Distribution Across Categories

The third key insight derived from the analysis pertains to the distribution of responses among the categories. In the CBID, out of 48 responses, 41 respondents chose "Academic Experience," 3 selected "Facilities," and four opted for "Safety, Well-being, and Inclusivity." Conversely, in the IBID, out of 30 responses, the distribution was different: 16 responses were related to "Safety, Well-being, and Inclusivity," 10 to "Academic Experience," 4 to "Facilities," and four responses did not fit into any of these categories and were classified as "Others." This comparison is shown in *Figure 6.1*

A chi-square test was performed to determine if the differences in category distribution between the two designs were statistically significant. The results indicated a significant difference in the distribution of responses across categories ($\chi^2 = 28.63$, $p < 0.001$). This suggests that the predefined categories in the CBID influenced respondents' choices, likely guiding them toward specific topics and limiting the scope of their feedback. As a result, the structured format may have constrained respondents, potentially skewing the data toward the explicitly provided categories.

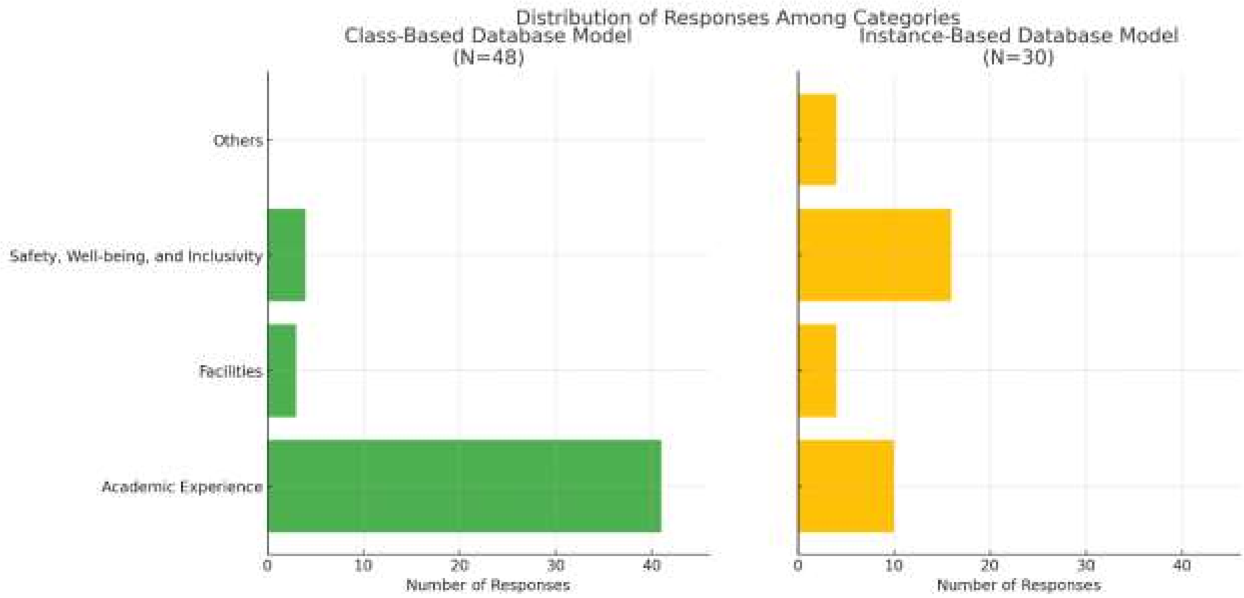

Figure 6.1 Comparative Response Distribution Across Categories

Another possible explanation for this distribution could be the primacy effect in the Class-Based condition. The first prompt in the dropdown list of the interface was "Academic Experience," which might have influenced respondents to select this option more frequently due to its prominent placement. This ordering could have unintentionally guided responses, highlighting the potential impact of prompt sequencing on data collection.

In contrast, IBID allowed respondents to freely express their insights without the constraint of predefined categories. This freedom resulted in a higher number of responses focusing on "Safety, Well-being, and Inclusivity," suggesting that when given the opportunity, respondents prioritized these issues over "Academic Experience" and "Facilities." This indicates that users had a greater interest or concern in safety and well-being topics, which may not have been fully captured in the structured CBID.

Additionally, the presence of responses categorized as "Others" in the IBID underscores the limitation of predefined categories. Despite thorough preparation, including pilot interviews and reviewing other surveys to determine appropriate categories, respondents still provided unique feedback that did not fit into the predetermined categories. This finding emphasizes the need for flexibility in survey design to capture a wider range of insights and perspectives.

Overall, this comparison reveals that predefined categories in the Class-Based design can influence and potentially limit the scope of respondents' feedback, whereas IBID's open-ended approach allows for a broader and potentially more genuine expression of respondents' concerns and priorities.

Upon analyzing the IBID responses, I obtained several key insights regarding their relevance, categorization, and depth of information provided in comparison to the CBID.

6.5. Effectiveness in Capturing Unanticipated Data

Of the 30 IBID responses, 28 were relevant to student life at MUN, while two were not. This high relevance rate (93.3%) indicates that the IBID effectively captures pertinent information from respondents.

The 28 relevant responses were further categorized into three main categories corresponding to those used in the CBID. Additionally, four responses were categorized under two different categories, indicating the presence of multifaceted insights in those responses. This suggests that the IBID allows for a more encompassing report of student experiences, as some responses provided information spanning multiple aspects of student life.

Additionally, when comparing the IBID responses to the predefined questions in the CBID, I found that 12 responses were included in CBID questions, and 16 responses were not. This comparison shows that more than half of the IBID responses (57%) provided insights that were not covered by the CBID questions. This demonstrates the ability of the IBID to uncover additional information and perspectives that a predefined questionnaire might overlook.

Further, of the 12 responses aligned with the CBID questions, 8 provided more detailed insights beyond the questions in the CBID, while 4 responses did not. This indicates that even when the IBID responses addressed topics included in the CBID, a significant majority (66.7%) offered more detailed, personalized, and unique experiences. This highlights the advantage of the IBID in capturing richer and more elaborate data.

In short, from these findings, these key implications for research can be derived:

Richness and Depth of Data: The IBID is superior in eliciting detailed and multifaceted insights from respondents. This is evidenced by the number of responses that provided additional information not covered by the CBID questions and those that offered more detailed personal experiences.

Flexibility and Scope: The IBID allows respondents to express a wider range of experiences and insights, which can lead to the identification of new themes and issues that were not initially considered in the survey design. This flexibility is particularly valuable for exploratory research, where the goal is to uncover a broad spectrum of perspectives.

Enhanced Understanding: The ability of the IBID responses to cover multiple categories suggests a more holistic understanding of student life. This could be especially useful for comprehensive assessments, where understanding the interplay between different aspects of student experiences is crucial.

Complementary Use: While the CBID is efficient for capturing specific, predefined information, the IBID complements it by providing depth and uncovering additional insights. A mixed-method approach that combines both designs could leverage the strengths of each, providing a more robust and comprehensive data collection strategy. A dynamic mixed design would optimize data collection by incorporating the benefits of both interface designs. In this approach, text boxes could be shortened to avoid discouraging respondents from providing feedback while avoiding predefined categories to prevent the introduction of biases.

One possible implementation of this design is to use AI capabilities to assist respondents dynamically. For example, the system could offer suggestions only after the respondent has expressed their initial thoughts, minimizing the risk of influencing their responses. This adaptive method would maintain the openness of the IBID while benefiting from the structured guidance of the CBID, ultimately capturing a broader and more authentic range of insights.

6.6. Ad hoc Analysis

6.6.1. Sentiment Analysis of IBID Responses

The sentiment analysis of the IBID responses revealed important insights into the effectiveness of the IBID compared to the CBID. Of the 30 IBID responses, 11 were positive, 11 were negative, three were no value-added, and 5 were improvement notes. In contrast, the CBID, which relied on predefined questions, yielded only positive and negative statements, with no improvement notes or no-value-added responses.

While the CBID included a text box at the end of each category for students to provide additional comments, only 8 out of 48 respondents chose to leave further feedback. Of these, one was a no-value-added statement; the remaining seven were negative. This contrast highlights the key advantage of the IBID: its open-ended nature encouraged more respondents to contribute additional, valuable feedback, including constructive improvement suggestions.

The Improvement Notes collected through the IBID provided particularly valuable insights, as they offered specific suggestions on how to improve various aspects of student life. This type of feedback was not possible to capture through the structured CBID, which restricted responses to more general sentiments. The presence of Improvement Notes in the IBID feedback offers actionable insights and a deeper understanding of student perspectives on what should change or be improved.

The IBID's flexibility acts as a double-edged sword. On the positive side, the IBID captured Improvement Notes, which were highly valuable for providing actionable insights and understanding what students believe should change. On the downside, it also captured No Value-Added Statements, which provided little relevant information. Nevertheless, the value of the constructive improvement notes outweighed the downside of receiving some non-informative responses.

For example, one respondent highlighted challenges faced by students in a professional program related to insufficient information provided ahead of time regarding course scheduling and required in-person components. They suggested that receiving a full program outline at the start would allow students more time to arrange necessary logistics such as accommodations, work leave, and childcare. This level of detailed, actionable feedback illustrates the kind of improvement-oriented insights that the IBID could capture, which the CBID likely would not have elicited.

As another example, one respondent noted that they had not participated in student activities or events due to personal circumstances and a busy schedule. While this feedback provides some context, it does not offer actionable insights or address specific issues related to the student experience at Memorial University. Such responses, though non-informative, were minimal in the overall dataset and did not detract from the valuable input captured through other responses.

The flexibility of the IBID allowed respondents to express their thoughts more freely and in greater detail, offering richer feedback compared to the CBID. This broader feedback provided deeper insights into the student experience, helping to identify aspects of university life that might be overlooked by the rigid structure of the CBID. While the occasional presence of no-value-added responses is a limitation, the IBID's ability to capture actionable and detailed feedback makes it a valuable tool for understanding student experiences and informing improvements.

In conclusion, the sentiment analysis demonstrates that the IBID is more effective in capturing a diverse range of feedback compared to the CBID. Although the IBID may yield some non-informative responses, its ability to gather improvement notes offers critical insights that the CBID cannot. The broader scope of feedback provided by the IBID, including personalized positive and negative statements and actionable suggestions, makes it a more powerful tool for understanding student experiences and identifying areas for improvement.

6.6.2. Comparative Analysis of AI Report Category

In the initial data collection using CBID, I predefined categories—Safety, Well-being, and Inclusivity; Facilities; and Academic Experience—and gathered responses accordingly. This approach aimed to streamline data into specific areas of interest. However, the majority of responses (85.4%) concentrated on Academic Experience, suggesting that the predefined categories may have limited the scope of feedback and overlooked other significant areas. Conversely, the IBID allowed for a more flexible and organic categorization of responses based on the issues mentioned by participants. This approach revealed a broader and more nuanced set of themes, such as Academic Experience, Student Life, Housing and Residence, Administrative Issues, Support Services, Financial Concerns, and the Impact of COVID-19. The IBID thus demonstrated that students' concerns were more diverse and multifaceted than initially captured by the predefined categories, underscoring the importance of an adaptable approach in accurately understanding and addressing the full range of student experiences.

Chapter 7

7. Limitations and Future Research

This research contributes to the broader discourse on interface design by highlighting the potential of IBID to revolutionize data collection practices, particularly in the context of UGC. The study underscores the importance of flexibility and depth in data capture, providing empirical evidence that supports the adoption of Instance-Based approaches in diverse and dynamic environments. Furthermore, this thesis extends the understanding of how different interface designs impact user engagement, offering practical insights for researchers and practitioners in the field. By demonstrating the advantages of IBID in capturing rich, detailed information, this work paves the way for future research and development in Instance-Based data management.

While this study provides valuable insights, it also has several limitations that should be addressed in future research. The study was conducted with a relatively small and homogenous sample, which may limit the generalizability of the findings. Future research should involve larger and more diverse populations to validate the results. The research focused on student life reporting, which may not fully capture the applicability of IBID in other contexts. Further studies should explore the use of IBID in the context of data collection in different domains, such as healthcare, marketing, and personalized services, to understand its broader applicability. Longitudinal studies could provide deeper insights into the long-term effects of using IBDM, particularly regarding data evolution and user engagement over time. This would help in understanding how well IBDMs can adapt to changing data landscapes.

Another limitation of the research was that each respondent could only fill out the survey once. This limitation was due to constraints in implementing the design. I am aware that allowing respondents to participate multiple times could have enriched the research by providing a more comprehensive understanding of user interactions with the interface designs over time. Future studies should consider incorporating mechanisms to enable multiple submissions, thereby capturing a wider range of data and insights.

Chapter 8

8. Conclusion

This thesis investigates the efficacy of IBID in capturing and representing UGC compared to traditional Class-Based designs. The analysis focused on the interface designs' ability to handle dynamic and diverse data inputs, emphasizing user engagement, data quality, and the richness of the information captured.

The study revealed that IBID offers superior flexibility, allowing respondents to provide more detailed and nuanced information. This design captured a broader spectrum of data, accommodating unique and unexpected insights often missed by the CBID. However, with its structured approach, the CBID yielded higher response and completion rates. This suggests that while IBID offers greater depth of information, their open-ended nature may be perceived as more demanding by respondents, potentially leading to higher abandonment rates.

Respondents using the IBID spent significantly more time completing the survey, reflecting the effort required to provide detailed, free-form responses. Conversely, the CBID facilitated quicker completion, which could be advantageous in large-scale surveys or contexts where brevity is essential. Qualitative analysis showed that the IBID responses were richer and more comprehensive, often including insights that extended beyond the predefined categories of the CBID. This highlights the value of IBID in capturing detailed and multi-faceted user experiences.

The distribution of responses across categories differed significantly between the two designs. The CBID's predefined categories appeared to limit the scope of feedback, while the IBID allowed respondents to focus on a broader range of issues, indicating a more genuine expression of their concerns and priorities.

In testing the hypotheses, evidence supported both Hypothesis 1 (H1) and Hypothesis 2 (H2). The following sections systematically examine each hypothesis, outlining the specific analyses that substantiate these claims and discussing their contributions to the overall evaluation of the hypotheses.

Hypothesis 1 (H1) posited that IBID would capture insights and information that CBID would not, particularly by accommodating the variability and uniqueness of UGC. This claim was strongly supported by multiple analyses.

First, the Effectiveness in Capturing Unanticipated Data (Section 6.5) demonstrated that 57% of IBID responses contained insights not covered by CBID’s predefined categories, confirming that CBID’s structured format constrained user input, whereas IBID facilitated the discovery of novel and unexpected insights.

Similarly, the Response Distribution Analysis (Section 6.4) revealed that 85.4% of CBID responses were concentrated in "Academic Experience," while IBID responses were more evenly distributed across multiple categories, indicating that CBID’s predefined categories biased user selection, limiting the diversity of responses. This further confirms IBID’s ability to capture a broader range of user-generated content.

Comparative Analysis of AI Report Categories (Section 6.7 – Ad Hoc Analysis)

Additionally, the Comparative Analysis of AI Report Categories (Section 6.7), though exploratory, provided further evidence that IBID captured broader and more varied themes, including topics such as housing concerns, financial struggles, and administrative challenges, which CBID failed to account for.

Hypothesis 2 (H2) asserted that CBID constrains user expression by imposing predefined categories, whereas IBID enables more authentic and unconstrained reporting. This hypothesis was supported by several key analyses.

The Average Word Count Analysis (Section 6.3) found that IBID responses contained an average of 55.9 words, more than twice the length of CBID responses, confirming that IBID allows users to articulate their experiences more freely and in greater detail.

Furthermore, the Survey Duration Comparison (Section 6.2) showed that IBID responses took an average of 430.03 seconds to complete, compared to 204.5 seconds for CBID, suggesting that IBID engages users in deeper, more thoughtful reflection, reinforcing the claim that it supports more authentic user expression.

Additionally, while not explicitly tied to H2, the Sentiment Analysis (Section 6.6) provided further evidence that IBID responses were not limited to binary positive or negative sentiments, as was the case with CBID. Instead, IBID responses included constructive improvement suggestions, highlighting that users felt more empowered to express nuanced opinions when unrestricted by predefined response formats. These findings substantiate H2, demonstrating that IBID fosters greater depth, authenticity, and expressive freedom in user responses.

The findings from this research have several implications for the design and implementation of interface systems, particularly in environments dealing with dynamic and diverse data inputs like UGC platforms. Given the strengths and weaknesses of both IBID and CBIDs, a hybrid approach may offer the best of both worlds. By integrating the structured framework of CBIDs with the flexibility of IBID, databases can provide both comprehensive data capture and ease of use. Enhancing the user experience is crucial as the higher abandonment rates associated with IBIDs suggest the need for improved user interface designs that guide respondents while still allowing flexibility. Techniques such as providing optional structured prompts within a free-form input framework could help balance user engagement with data richness.

The choice of an interface design should be tailored to the specific needs of the application. For scenarios where detailed qualitative data is crucial, IBID would be more appropriate. Conversely, CBIDs may be preferable for applications requiring high response rates and standardized data. Implementing IBID may require additional training and support for users to ensure they understand how to effectively provide detailed responses. This can help mitigate the challenges associated with the open-ended nature of IBID and enhance data quality.

This study contributes to the literature by providing empirical evidence on the impact of interface design on UGC collection, addressing gaps in the study of Class-Based and Instance-Based interface designs. It extends prior work in conceptual modeling by validating the effectiveness of IBID in a new domain: student life reporting, demonstrating its advantages in capturing unstructured, detailed, and nuanced user experiences. Additionally, the findings contribute to practice by offering actionable insights for institutions and practitioners designing UGC reporting systems, reinforcing the necessity of hybrid interface models that optimize both structured guidance and user flexibility.

In conclusion, this thesis has demonstrated the potential of IBID to capture rich, detailed, and nuanced UGC, offering significant advantages over traditional CBID in dynamic and diverse data environments. While IBID presents certain challenges in terms of user engagement and survey completion rates, its ability to uncover unanticipated insights and provide comprehensive data makes it a valuable tool in the modern data management landscape. Future research and technological advancements will likely continue to refine and enhance the application of IBID, contributing to more effective and insightful data management practices across various domains.

Bibliography

1. Abrahão, S., Insfran, E., Sluÿ7fters, A., & Vanderdonckt, J. (2021). Model-based intelligent user interface adaptation: Challenges and future directions. *Software and Systems Modeling, 20*, 1-23. https://doi.org/10.1007/s10270-021-00909-7

2. Araújo, C., Oliveira, M., Nogueira, B., Maciel, P., & Tavares, E. (2024). Performability evaluation of NoSQL-based storage systems. *The Journal of Systems and Software, 208*, 111885. https://doi.org/10.1016/j.jss.2023.111885

3. Bagui, S. (2003). Achievements and weaknesses of object-oriented databases. *Journal of Object Technology, 2*(4), 29-41. https://doi.org/10.5381/jot.2003.2.4.c2

4. Bączek, M., Zagańczyk-Bączek, M., Szpringer, M., Jaroszyński, A., & Wożakowska-Kapłon, B. (2021). Students' perception of online learning during the COVID-19 pandemic: A survey study of Polish medical students. *Medicine, 100*(7), e24821. https://doi.org/10.1097/MD.0000000000024821

5. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

6. Brenner, M. (1985). *The research interview: Uses and approaches*. Academic Press.

7. Brodie, M. L., Mylopoulos, J., & Schmidt, J. W. (1984). On the development of data models. In *On conceptual modelling* (pp. 19-47). Springer New York. https://doi.org/10.1007/978-1-4612-5196-5_2

8. Brown, A. W. (1991). *Object-oriented databases: Applications in software engineering*. McGraw-Hill.

9. Cattell, R. (2011). Scalable SQL and NoSQL data stores. *ACM SIGMOD Record, 39*(4), 12-27.

10. Chen, P. P. S. (1975). The entity-relationship model: Toward a unified view of data. *SIGIR Forum, 10*(3), 9-9. https://doi.org/10.1145/1095277.1095279

11. Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of the ACM, 13*(6), 377-387.

12. DBTG Codasyl. (1971). CODASYL data base task group report. In *Conference on Data Systems Languages*. ACM.

13. Dean Meltz, R. L., Harrington, M., Hain, R., & Nicholls, G. (2004). *An introduction to IMS: Your complete guide to IBM's Information Management System*. IBM Press.

14. DeWitt, D. J., & Levine, C. (2008). Not just correct, but correct and fast: A look at one of Jim Gray's contributions to database system performance. *SIGMOD Record, 37*(2), 45-49.

15. Embley, D. W., & Thalheim, B. (2011). *Handbook of conceptual modeling: Theory, practice, and research challenges*. Springer. https://doi.org/10.1007/978-3-642-15865-0

16. Fettke, P. (2009). How conceptual modeling is used. *Communications of the Association for Information Systems, 25*, 571-592.

17. Grolinger, K., Higashino, W. A., Tiwari, A., & Capretz, M. A. M. (2013). Data management in cloud environments: NoSQL and NewSQL data stores. *Journal of Cloud Computing: Advances, Systems and Applications, 2*(1), 22.

18. Lieder, F., Griffiths, T. L., Huys, Q. J. M., & Goodman, N. D. (2018). The anchoring bias reflects the rational use of cognitive resources. *Psychonomic Bulletin & Review, 25*(1), 322-349. https://doi.org/10.3758/s13423-017-1286-8

19. Lukyanenko, R., Parsons, J., & Samuel, B. M. (2018). Representing instances: The case for reengineering conceptual modelling grammars. *European Journal of Information Systems, 28*(1), 68-90. https://doi.org/10.1080/0960085X.2018.1488567

20. Lukyanenko, R., Wiersma, Y. F., Huber, B., Parsons, J., Wachinger, G., & Meldt, R. (2017). Representing crowd knowledge: Guidelines for conceptual modeling of user-generated content. *Journal of the Association for Information Systems, 18*(4), 297-339. https://doi.org/10.17705/1jais.00456

21. Parsons, J. (2003). Data modeling. In Błażewicz, J., Kubiak, W., Morzy, T., & Rusinkiewicz, M. (Eds.), *Handbook on data management in information systems* (pp. 77-94). Springer. https://doi.org/10.1007/978-3-540-24742-5_3

22. Parsons, J., & Wand, Y. (2000). Emancipating instances from the tyranny of classes in information modeling. *ACM Transactions on Database Systems (TODS), 25*(2), 228-268.

23. Pedersen, T., Johansen, C., & Jøsang, A. (2018). Behavioural computer science: An agenda for combining modelling of human and system behaviours. *Human-Centric Computing and Information Sciences, 8*(7). https://doi.org/10.1186/s13673-018-0130-0

24. Shanks, G., Tansley, E., Nuredini, J., Tobin, D., & Weber, R. (2008). Representing part-whole relations in conceptual modeling: An empirical evaluation. *MIS Quarterly, 32*(3), 553-573. https://doi.org/10.2307/25148856

25. Silberschatz, A., Korth, H., & Sudarshan, S. (1996). Data models. *ACM Computing Surveys, 28*(1), 105-108.

26. Stanford University. (n.d.). *Dean of students - Student of concern form*. Stanford University.

27. Stonebraker, M., & Rowe, L. (1990). Third-generation database system manifesto. *Memorandum No. UCB/ELB. M90/23*. University of California, Berkeley.

28. Taipalus, T. (2024). Database management system performance comparisons: A systematic literature review. *The Journal of Systems and Software, 208*, 111872. https://doi.org/10.1016/j.jss.2023.111872

29. Watt, A., & Eng, N. (2014). *Database design*. BC Campus, BC Open Textbook Project.

30. Won Kim. (1990). Object-oriented databases: Definition and research directions. *IEEE Transactions on Knowledge and Data Engineering, 2*(3), 327-341.

Appendix A

A CBID Reference Mapping

The table below provides an overview of how each question in the survey was inspired by specific reference sources. The table includes a list of all survey questions along with corresponding references that informed the development of each question. Where a reference contributed to the creation of a particular question, it is marked with a check.

Table A.1Mapping of Survey Questions to Reference Sources

	References				
CBID Question	ISOC (2024) – Question Wording Templates	Stanford University – Dean of Students Form	UT Austin (2023) – Instructor Report, AAS S312	Bączek et al. (2021) – COVID-19 Online Learning Survey	University of Toronto (2020) – ISA Report
Please indicate your academic degree:	X				
Please indicate your academic program:	X				
In what year of your program are you?	X				
What would you like to report on?		X			
Rate the instructional effectiveness of the courses.			X	X	X
Rate how organized the course materials were.			X	X	X
Rate the clarity of instruction provided by the course instructors.			X		
Rate the availability of the instructors for office hours and additional help.			X	X	
Rate the level of engagement in the classes.			X	X	X
Rate the quality of the lectures.			X	X	
Rate the quality and relevance of the course assignments.					X
Rate the fairness of grading in courses.			X		
Rate how available the instructors were for providing feedback.			X		X

Rate how effective the course advising was in preparing you for classes.					X
Rate the usefulness of the course materials.			X		
Rate how likely you are to recommend these courses to another fellow student.					X
Rate how helpful the Teacher's Assistants were.					X
Select the campus facility you are providing feedback on:	X				
Rate the level of cleanliness of the facility.					X
Rate the quality of service received at the facility.					X
Rate the accessibility of the facility for differently-abled individuals.					X
Rate your satisfaction with the helpfulness of the staff.					X
Rate your satisfaction with the operating hours of the facility.					X
Rate your overall satisfaction with the facility.					X
Rate the responsiveness level of the staff in addressing issues.		X			X
Rate the visibility and responsiveness of security personnel in bias-related incidents.					X
Rate how well the campus communicates available safety, well-being, and inclusivity resources.					X
Rate your confidence in the university's ability to effectively handle incidents of bias or discrimination.					X
Rate the effectiveness of the aspect of bias reporting.		X			
Rate your perception of inclusion and acceptance on campus.					X
Rate your sense of physical safety on campus.		X			X
Rate your level of comfort when interacting with staff and faculty.		X			X

Appendix B

B Survey Customization Based on Interview Feedback

The table below provides three examples of how feedback from initial interviews influenced the development and refinement of the survey questions. Each example highlights a key issue raised during the interviews, the corresponding survey question that was inspired by the feedback, and the adjustments made to the final version of the question. It is noteworthy to add that in order to protect the anonymity of the interviewees; I have chosen to provide a summary of their responses rather than their exact words. This approach ensures confidentiality while still conveying the key insights and feedback gathered during the interviews.

Table B.1 Customization Based on Interview Feedback

Issue Raised	Category	Survey Questions Inspired	Adjustments Made
The student voiced frustration over the quality and cost of food on campus, as well as the accessibility and cleanliness of campus facilities.	Facilities	Rate the level of cleanliness of the facility.' 'Rate the quality of service received at the facility.' 'Rate your satisfaction with the operating hours of the facility.'	The student's comments emphasized the need to ensure that food and facility-related questions were included and clearly framed to capture these concerns. The feedback helped prioritize cleanliness and service-related questions, which were seen as critical to the overall student experience.
The student mentioned witnessing racism and acknowledged the university's efforts in promoting inclusivity, but also touched upon issues	Safety, Well-being, and Inclusivity:	Rate your perception of inclusion and acceptance on campus.' 'Rate your sense of physical safety on campus.' 'Rate how well the campus communicates available safety, well-being, and inclusivity resources.'	The student's comments on racism and safety led to additional refinement of questions related to campus inclusivity and safety. The survey was adjusted to capture perceptions of inclusivity, as well as the effectiveness of safety resources and communication.

outside the campus that impacted safety.			
The student highlighted problems with the lack of structure in the Computer Science department, poor performance of Teaching Assistants (TAs), and the inconsistency in the quality of professors.	Academic Experience:	'Rate the instructional effectiveness of the courses.' 'Rate how helpful the Teacher's Assistants were.' 'Rate the clarity of instruction provided by the course instructors.'	The student's input reinforced the need to emphasize the role of TAs and the structure of academic experiences. The wording of the questions was adjusted to reflect specific concerns about instructional effectiveness and support from teaching staff.

Appendix C

C CBID

Entities and Attributes:

- Report
  - Report ID(PK)
  - Academic Program
  - Academic degree
  - Year of Program
  - Issue category

- Academic Experience
  - Academic Experience ID
  - Instructional Effectiveness Satisfaction Rating
  - Course Material Organization Satisfaction Rating
  - Clarity of Instruction Satisfaction Rating
  - Instructor Availability Satisfaction Rating
  - Class Engagement Satisfaction Rating
  - Lecture Quality Satisfaction Rating
  - Course Assignments Satisfaction Rating
  - Grading Fairness Satisfaction Rating
  - Instructor Feedback Availability Satisfaction Rating
  - Course Advising Effectiveness Satisfaction Rating
  - Course Material Usefulness Satisfaction Rating
  - Course Recommendation Likelihood Satisfaction Rating
  - TA Helpfulness Satisfaction Rating
  - Text report
  - Report ID (FK)

- Facilities
  - Facilities ID (PK)
  - Facility Name (e.g., Library, Gym, Food Court, Student Center)

  - Facility Cleanliness Satisfaction Rating
  - Service Quality Satisfaction Rating
  - Facility Accessibility Satisfaction Rating
  - Staff Helpfulness Satisfaction Rating
  - Operating Hours Satisfaction Rating
  - Overall Satisfaction Rating
  - Staff Responsiveness Satisfaction Rating
  - Text report
  - Report ID (FK)

- Safety, Well-being, and Inclusivity

  - Safety ID (PK)
  - Security Personnel Responsiveness Satisfaction Rating
  - Available Resources Campus Communication Satisfaction Rating
  - Bias Incident Handling Confidence Satisfaction Rating
  - Bias Reporting Effectiveness Satisfaction Rating
  - Campus Inclusion Perception Satisfaction Rating
  - Sense of Physical Safety Satisfaction Rating
  - Staff/Faculty Interaction Comfort Satisfaction Rating
  - Text report
  - Report ID (FK)

- Relationships:

  - Report to Academic Experience: One-to-zero or one
  - Report to Facilities: One-to-zero or one
  - Report to Safety, Well-being, and Inclusivity: One-to-zero or one

The ERD explained above is depicted below.

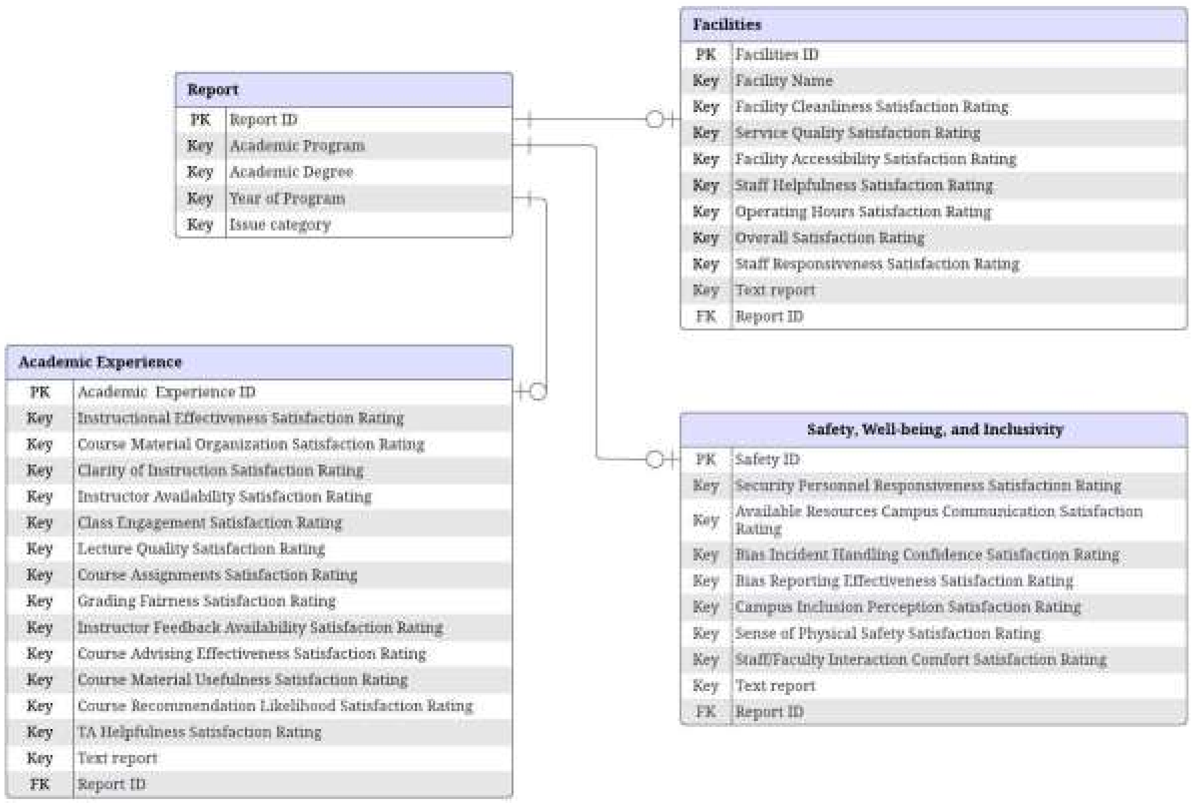

Figure C.1 Class-Based ERD

Appendix D

D University Student Experience Survey

University Student Experience Survey

The survey content is structured as follows: The student will be shown the consent form that is attached to the application. After clicking on the "Accept" button, they will begin the survey. First, I will gather demographic information, and then students will be randomly assigned to either an Instance-Based or Class-Based design. In the CBID, there are three categories from which respondents can choose to report. Depending on their selection, they will be presented with a corresponding block of questions. Upon completing the survey, participants will have the option to provide their email address. After this, a completion message will be displayed.

The content is provided below:

Demographic Information

Please indicate your academic degree:

- Bachelor's degree
- Masters degree
- Doctoral degree
- Diploma/Certificate
- Graduate Diploma
- Post-Graduate/Resident

Please indicate your academic program:

- Arctic and Sub-Arctic Studies (Labrador

- Institute)
- Business Administration
- Education
- Engineering and Applied Science
- Human Kinetics and Recreation
- Humanities and Social Sciences
- Medicine
- Music
- Nursing
- Pharmacy
- Science
- Social Work
- Arts and Social Science (Grenfell Campus)
- Fine Arts (Grenfell Campus)
- Science and the Environment (Grenfell Campus)
- Fisheries (Fisheries and Marine Institute)
- Maritime Studies (Fisheries and Marine
- Institute)
- Ocean Technology (Fisheries and Marine
- Institute)
- Other: ____

In what year of your program are you?

- Year 1
- Year 2
- Year 3
- Year 4
- Year 5 or beyond

Class-Based Survey Questions

What would you like to report on?

- Academic Experience

- Facilities
- Safety, Well-being, and Inclusivity

Academic Experience

Please rate your satisfaction with the following academic services during this semester on a scale of 1 to 5 (5 being the highest and 1 being the lowest):

- Rate the instructional effectiveness of the courses.
- Rate how organized the course materials were.
- Rate the clarity of instruction provided by the course instructors.
- Rate the availability of the instructors for office hours and additional help.
- Rate the level of engagement in the classes.
- Rate the quality of the lectures.
- Rate the quality and relevance of the course assignments.
- Rate the fairness of grading in courses.
- Rate how available the instructors were for providing feedback.
- Rate how effective the course advising was in preparing you for classes.
- Rate the usefulness of the course materials.
- Rate how likely you are to recommend these courses to another fellow student.
- Rate how helpful the Teacher's Assistants were.

Please provide any further information you wish to provide regarding your report:

____

Facilities

Select the campus facility you are providing feedback on:

- St. John's Campus

- Grenfell Campus
- Labrador institute
- Marine Institute
- Harlow Campus
- Burton's Pond Apartments
- Macpherson College
- Paton College
- Signal Hill Campus
- Other

Please rate your satisfaction with the following on a scale of 1 to 5 (5 being the highest and 1 being the lowest):

- Rate the level of cleanliness of the facility.
- Rate the quality of service received at the facility.
- Rate the accessibility of the facility for differently-abled individuals.
- Rate your satisfaction with the helpfulness of the staff.
- Rate your satisfaction with the facility's operating hours.
- Rate your overall satisfaction with the facility.
- Rate the responsiveness level of the staff in addressing issues.

Please provide any further information you wish to provide regarding your report:

____

Safety, Well-being, and Inclusivity

Please rate your satisfaction with the following on a scale of 1 to 5 (5 being the highest and 1 being the lowest):

- Rate the visibility and responsiveness of security personnel in bias-related incidents.
- Rate how well the campus communicates available safety, well-being, and inclusivity resources.

- Rate your confidence in the university's ability to effectively handle incidents of bias or discrimination.
- Rate the effectiveness of the aspect of bias reporting.
- Rate your perception of inclusion and acceptance on campus.
- Rate your sense of physical safety on campus.
- Rate your level of comfort when interacting with staff and faculty.

Please provide any further information you wish to provide in regard to your report:

____

Instance-Based Survey Question

We invite you to share your thoughts and experiences regarding your student life at our university through this form. This is your opportunity to communicate anything that matters to you, whether it is a memorable experience, a particular challenge you've faced, suggestions for improvement, concerns about campus policies, or your perspectives on university initiatives. Your insights are invaluable in helping us understand the diverse aspects of student life and in fostering a supportive and dynamic university community.

____

____

Appendix E

E Distribution Tables

Table E.1 Distribution of Academic Degrees

Degree	***Number of Respondents***
Master's degree	28
Bachelor's degree	32
Graduate Diploma	4
Doctoral Degree	13
Post-Graduate/Resident	1

Table E.2 Distribution of Academic Programs

Program	***Number of Respondents***
Arts and Social Science (Grenfell Campus)	1
Business Administration	39
Education	3
Engineering and Applied Science	10
Humanities and Social Sciences	5
Maritime Studies (Fisheries and Marine)	1
Medicine	1
Ocean Technology (Fisheries and Marine)	1
Other	4
Science	11
Science and the Environment (Grenfell Campus)	1
Social Work	1

Table E.3 Distribution of Program Years

Program Year	***Number of Respondents***
Year 1	29
Year 2	20
Year 3	11
Year 4	11
Year 5 or beyond	7

Appendix F

F VBA Code for Wordcount function

```
Function WordCount(cell As Range) As Integer
    Dim text As String
    Dim wordArray() As String

    ' Get the text from the cell and trim leading and trailing spaces
    text = Trim(cell.Value)

    ' Handle empty cell
    If Len(text) = 0 Then
        WordCount = 0
        Exit Function
    End If

    ' Split the text by spaces into an array
    wordArray = Split(text, " ")

    ' Count the number of elements in the array
    WordCount = UBound(wordArray) - LBound(wordArray) + 1
End Function
```