CHATGPT: TODAY THE WORLD, TOMORROW EDUCATION?

UNDERSTANDING PERCEPTIONS OF CHATGPT AMONG ACADEMIC FACULTY AND STAFF

THOMAS BROLIN STJÄRNE

ALEXANDER FRIEL

Bachelor Thesis
Stockholm School of Economics
2024

ChatGPT: Today the world, tomorrow education?

Abstract:

This thesis investigates the perceptions of academic faculty and staff towards the use of ChatGPT within Swedish higher education institutions. With the Technology Acceptance Model (TAM) as the theoretical framework, this study explores the factors influencing the perceptions underlying acceptance and use of ChatGPT. Utilizing a quantitative approach with a survey methodology, data is collected from 75 participants across 24 institutions. Its findings reveal significant correlations between the perceived usefulness of ChatGPT and image, output quality, and result demonstrability, whereas perceived usefulness was determined to be a non-significant factor. Subjective norm and job relevance were also identified to have significant mediated effects through image. The study further concludes that these significant correlations are maintained both when measuring the perceived usefulness for faculty usage versus that of student usage of ChatGPT.

The results are intended to contribute to the discourse on AI implementation in higher education as well as assist educational administrators and policymakers in making informed decisions when considering strategies for integrating AI tools in teaching and administrative processes.

Keywords:

Artificial Intelligence (AI), Generative AI, ChatGPT, Technology Acceptance Model (TAM), Technology Implementation in Academia, Information Systems Discourse

Authors:

Thomas Brolin Stjärne (student 24591)
Alexander Friel (student 25455)

Supervisor:

Wijkström, Filip, Associate Professor, Department of Management and Organization

Examiner:

Laurence Romani, Professor, Department of Management and Organization

Bachelor Thesis
Bachelor Program in Management
Stockholm School of Economics

Contents

We would like to thank Laurence for constantly encouraging us to widen our ideas and perceptions, Filip for teaching us how narrow down and concretize these very same ideas and perceptions, and Rebecka for helping us find our way in the early stages of a long process.

"So long and thanks for all the fish" – Douglas Adams

1. Introduction

1.1. Background

In 2016 the World Economic Forum called artificial intelligence (AI) the Fourth Industrial Revolution, referencing its ability to take automation one step further. While previous computer programs required precise coding activities, AI systems rely on large databases and use classes of algorithms to map tasks in an autonomous way (Bação, Lopes, & Simões, 2023). With 80% of organizations seeing AI as a strategic opportunity and 85% viewing it as a way to achieve competitive advantage, many organizations are investing in AI technologies (Enholm et al., 2021). However, discourse within Information Systems (IS) research suggests that AI technology is fundamentally different from traditional information technology (IT), with organizations failing to realize their expected value. Thus from the perspective of strategic management, organizations must adequately prepare to drive effective AI implementation and capture its benefits (Weber et al., 2022).

The complexities of implementing AI technology are expected to extend beyond organizations and individuals who adopt them, driving industrial-scale and societal changes on a global level (Saghafian et al., 2021). When ChatGPT was released in 2022 it was the first generation of generative AI (GenAI) chatbots capable of generating output with a high enough degree of sophistication to cause academic institutions to question their internal procedures (Sullivan, Kelly & McLaughlan, 2023). Indirect technological changes caused by ChatGPT has therefore led to academic discussions regarding organizational changes within higher educational institutions. Organizational readiness is an organization's ability to adapt to required changes, with technological changes encompassing technological infrastructure, managerial commitment to change, knowledge and (human) resources (Saghafian et al., 2021).

However, ChatGPT does not only present a threat to academia, it also has the potential to facilitate learning and increase administrative efficiency. Despite the potential benefits that the implementation of GenAI systems could provide, its adoption within the field of education has been slow compared to other sectors such as industry and medicine (Kasneci et al., 2023). This indicates a need to observe the factors underlying the adoption, in order to evaluate the degree of organizational readiness in higher education.

1.2. Research gap

As recent IS research suggests that AI technology is fundamentally different from traditional IT, there is a large research gap bridging between the traditional tested models and theories, and these technologies (Weber et al., 2022).

Our research gap was identified by comparing two systematic reviews. The first article by Crompton and Burke (2023) reviewed 138 articles in the field of AI in higher education between the years of 2016 and 2022 and identified a predominant focus on students (72%) in AI research in higher education, with instructors (17%) and managers (11%) receiving less attention. The same review also indicated that there seemed to be a lack of empirical research on the subject. The second systematic review by Bahroun et al. (2023) reviewed 217 research papers published between 2018 and 2023 which highlighted a sharp rise in publications between 2022 and 2023. It found ChatGPT to be predominant mentioned AI tool, occurring in 151 papers, making it an appropriate measurement proxy.

1.3. Purpose and Research Question

The main purpose of this thesis is to contribute to the discourse in management and information systems research on how to drive and manage AI implementation in organizations. The study aims to investigate how perceptions of ChatGPT among academic faculty and staff influence their intention to implement its use in an academic setting. This will be accomplished by measuring and quantifying the factors influencing academic faculty and staff's perceptions of both faculty and student use of ChatGPT. Framed as a research question, we seek to answer:

What underlying factors hold significant explanatory value for the acceptance of Artificial Intelligence in Academia among academic faculty and staff?

1.4. Scope

This study has been geographically delimited to academic faculty and staff at Swedish universities or university colleges (Swe. *Högskolor*) for reason of managing contact within the time scope of the study. Data for this study was gathered from 24 institutions across seven schools. These institutions have a high variation in current implementation and policies regarding ChatGPT. As the definition of what encompasses *faculty* varies across institutions, this study leads with the definition of teaching staff. Respondents were given an opportunity to define their role as being either faculty, researcher, administrative staff or self-defined.

2. Literature review

2.1. Qualitative academic perceptions of Artificial Intelligence

A qualitative study by Firat (2023) explored the perspectives of students and educators on the implication of ChatGPT and AI integration in higher education. This study used a thematic content analysis and found that integrating AI in education offers opportunities to enhance learning experiences and transform the role of educators, but also that the shift entails challenges in assessment, digital literacy, and ethical considerations.

Similar sentiments were established in a text mining analysis of 86 articles published in scientific journals between 2022 and 2023, which showed an overall positive sentiment albeit with concerns of potential misuse and plagiarism (Bringula, 2023). Identified research has at the point of writing this literature review focused primarily on identifying academic opportunities and challenges surrounding ChatGPT, and not on underlying drivers of perception.

2.2. Technology acceptance in Higher education

Measuring user acceptance is a way of determining a teacher's attitude toward using new technologies in their educational practices. Two well established theories describe the mechanisms and factors affecting technology adoption; the Unified Theory of Acceptance and Use of Technology (UTAUT) and the Technology Acceptance Theory (TAM) (Scherer, Siddiq, and Tondeur, 2019).

Granić (2023) finds that TAM and UTAUT are considered the two most prominent theoretical approaches in researching technology acceptance in educational contexts. UTAUT has been shown to be hold good explanatory ability but is primarily directed at consumer behavior analysis and has received criticism for including too many independent constructs for predicting intentions and behavior. TAM on the other hand is the most widely used validated model for prediction and explanation of user's behavior toward acceptance and adoption of educational technology. The model is known for its

parsimony, or minimization of number of parameters, which is a strength for practical, predictive applications. Informed by this, the technology acceptance model has been chosen as the model on which our research question is tested.

3. Theoretical framework

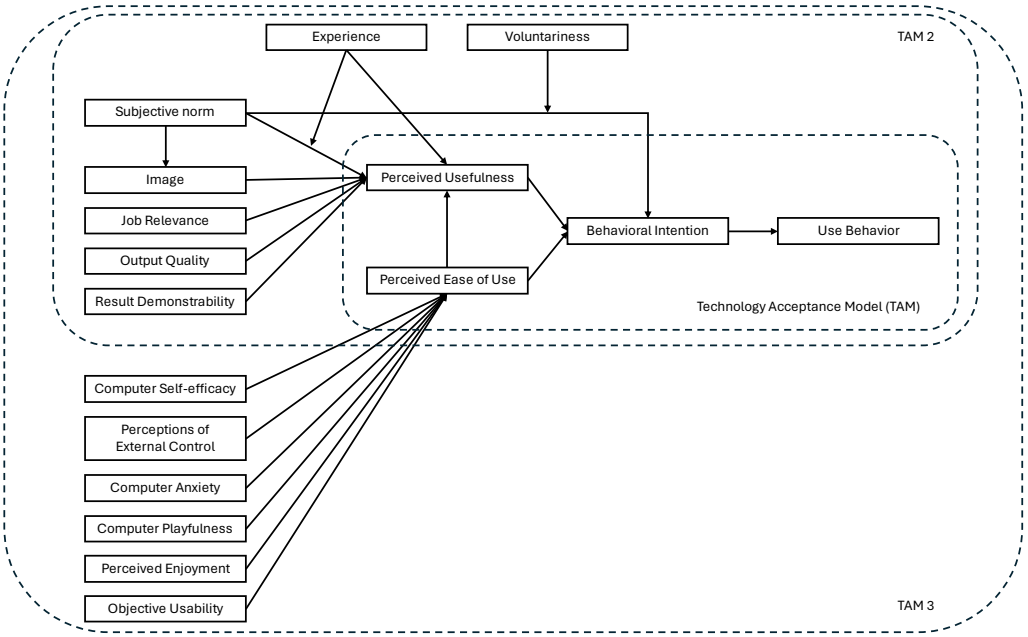
3.1. Technology Acceptance Model, an overview

The Technology Acceptance Model, first proposed by Davis (1989), and its further iterations TAM 2 (Venkatesh & Davis, 2000) and TAM 3 (Venkatesh & Bala, 2008), are models that seek to delineate the factors influencing users' acceptance and usage of technology. The original model established *perceived usefulness* and *perceived ease of use* as the primary determinants of technology acceptance via informing *behavioral intent* that then translates into *use behavior*.

TAM 2 extended the model by introducing cognitive instrumental processes (*job relevance*, *output quality*, and *result demonstrability*) and social influence processes (*subjective norm*, *voluntariness*, and *image*), which together with perceived ease of use act as antecedents to perceived usefulness, the latter two in turn significantly influencing user acceptance. TAM 2 emphasizes the role of technology acceptance in an organizational context (Venkatesh & Davis, 2000).

TAM 3 further broadens the framework by integrating additional determinants for perceived ease of use, such as computer self-efficacy, perceptions of external control, computer anxiety, and computer playfulness (Venkatesh & Bala, 2008).

Figure 1: Technology Acceptance Model, iterations 1, 2 and 3.



3.2. TAM applicability in an academic context

TAM and its extended models have seen a widespread adoption in research on technology acceptance in the educational field. A meta-study of 71 primary studies (Granić & Marangunić, 2019) found that there is vast evidence supporting perceived ease of use and perceived usefulness as antecedent factors affecting acceptance of learning with technology and that the TAM is both a leading scientific paradigm and credible model for facilitating assessment of diverse technological deployments in educational context. A research gap identified by the same study was that most established research has been conducted with university students as the sample group, and that academics/teachers have seen limited direct study.

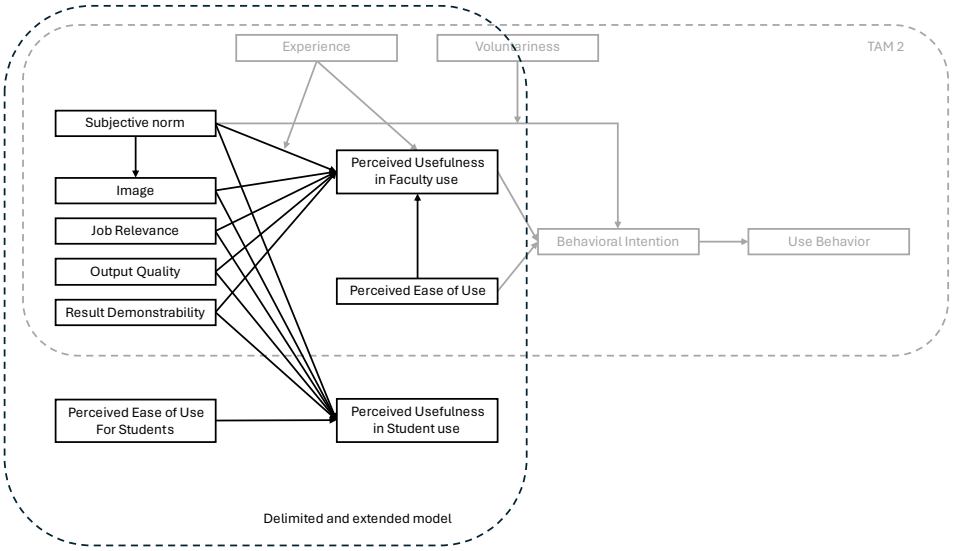
3.3. Choice and delimitation of model

Under the assumption that perceived ease of use would be less relevant among academics who conduct much of their work digitally, as well as the low complexity of ChatGPT's user interface, the TAM 2 model with its focus on the determinants of perceived usefulness was chosen to maximize research parsimony (seeking the simplest explanation that fits the evidence) and to emphasize the utility perspective.

An issue encountered in our preparatory process was the varying restrictions on usage of ChatGPT in place among the institutions where we sought to survey. This in conjunction with the early stage of ChatGPT prevalence meant that behavioral intent would be inherently hard to establish due to external biases.

Given the previously validated general relationship between perceived usefulness, perceived ease of use, and behavioral intent on a general level within academia (Granić, 2023), our research instead focuses on affirming the relationships between the determinants of perceived usefulness. We also hypothesized a potential difference in the view on utility provided by ChatGPT to faculty versus students. The result was a delimitation of the model as well as the introduction of two mirrored perceived usefulness dependent variables, one for faculty usage and one for student usage.

Figure 2: Delimitation and extension of TAM 2 as applied in study.



This informed the following hypotheses:

H1	Perceived Ease of Use has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use.
H2	Subjective Norm has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use
H3	Subjective Norm has an indirect, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use through mediation via Image.

H4	Image has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use.
H5	Job Relevance has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use.
H6	Output Quality has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use.
H7	Result Demonstrability has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use.
H8	Perceived Usefulness of ChatGPT in student use is significantly different within the sampled population compared to Perceived Usefulness of ChatGPT in faculty use.
H9	Perceived Usefulness of ChatGPT in student use shares the underlying significant independent variables that present for Perceived Usefulness of ChatGPT in faculty use.

4. Methodology

4.1. Research paradigm

Our research is grounded in objectivist ontology and positivist epistemology. This entails an underlying assumption of an objective reality that can be deduced through observation of phenomena (Saunders et al., 2019). Our goal is thus to either accept or falsify our hypotheses in order to generate statements about reality that are objectively true.

4.2. Research Design

We set out to conduct a cross-sectional quantitative study where we constructed a survey to be sent out to recipient faculties. The Technology Acceptance Model and its derivatives are well-established as the leading model within the field of technology acceptance, with over 50,000 citations in between the articles of TAM 1, 2, and 3. As our literature review did not uncover previous quantitative TAM research in the intersection of Faculty and ChatGPT use in Higher Education, our research design process has been informed by general research applying the Technology Acceptance Model framework.

4.2.1. Time Horizon

The original studies for the Technology Acceptance Model and its derivatives are longitudinal in nature, measuring the change in acceptance over time. Due to time constraints we elected to perform a cross-sectional study intent on informing the situation at current point in time.

4.2.2. Pilot Study

Due to said time constraint and our initially limited access to academic institutions, a full pilot study could not be conducted. In order to still evaluate our questionnaire, we established contact with a deputy Head of Department and three academics at a Swedish **et a**university institution who provided iterative feedback on the questionnaire design, serving as a non-random small-scale pilot group or panel. Input was also provided by our course supervisor and members of the Department of Management and Organization at the Stockholm School of Economics.

4.2.3. Research Ethics and GDPR

Throughout the design process we relied on the questionnaire design advice and ethical research guidelines outlined by Saunders et al. (2019) to structure both the survey and our approach to reaching faculty members and staff at academic institutions across Sweden.

Feedback during the design of our survey saw us alter the demographic section to further ensure no information could be tied back to an individual or even singular institution. Our chosen method of survey distribution through official channels meant we felt confident in avoiding the collection of metadata (i.e., IP addresses). This allowed us to gather a fully anonymous sample that retained no identifiable markers, ensuring full compliance with the General Data Protection Regulation through its article 25 on Data protection by design and by default (GDPR, 2016).

4.2.4. Questionnaire and Measure design

The phenomenon we wished to study was underlying determinants of the perceived usefulness of ChatGPT in use by faculty and students, among academic faculty and staff

at academic institutions. We adapted the TAM model for this purpose by delimiting our study to measure *Perceived Usefulness* (PU) (and its determinants) of ChatGPT in a higher educational context, following the TAM 2 model.

In order to capture views on ChatGPT as an academic tool for these two types of users, PU was further recontextualized into the two variables *Perceived Usefulness for Faculty* (PU-F) and *Perceived Usefulness for Students* (PU-S). In terms of determinants of Perceived Usefulness, all measures outlined in TAM 2 were included (albeit Experience, a longitudinal measure, was recontextualized). While we wanted to measure the individual's PEOU, we theorized that *Perceived Ease of Use for Students* (PEOU-S) could be significantly different from PEOU and thus introduced as a separate variable. A similar idea informed the decision to split *Result Demonstrability* (RES) into the separate variables *Result Demonstrability for Faculty* (RES-F) and *Result Demonstrability for Students* (RES-S), based on the assumption that the perceived sough result of ChatGPT use would be different depending on user type.

An additional extension to the model was made by subdividing the TAM2 measure *Experience* (EXP) into *Familiarity* and the Boolean control variable "have previously used ChatGPT". Experience had a longitudinal effect on Behavioral Intention and Subjective Norm in the original TAM 2 and TAM 3 studies (Venkatesh & Davis, 2000, Venkatesh & Bala, 2008) and this extension was made both to be able to test an experienced subsample and to control for effects of familiarity with ChatGPT on the dependent variables.

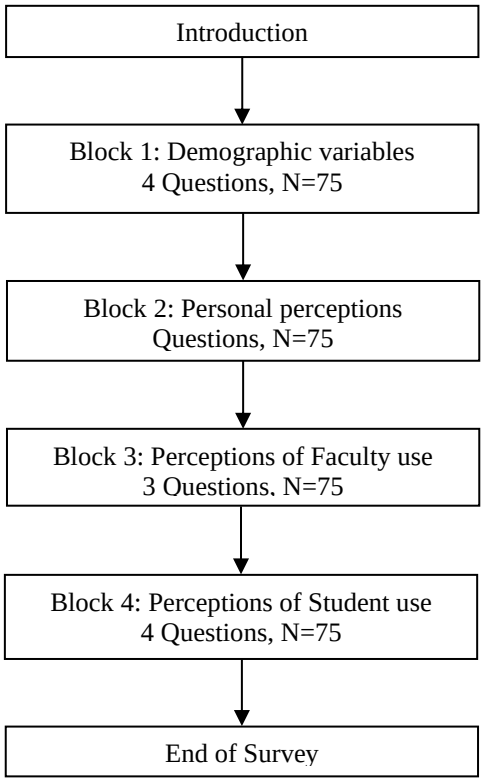
Early on in our thesis process our supervisors expressed worry over the availability and general questionnaire fatigue among our target population, with associated low expectations regarding response rate. With the small time-frame available to us, we decided to rely on single- or two-item measures rather than multi-item constructs, in order to promote participation through brevity. As a pilot study could not be deployed within the time frame, evaluation was conducted by allowing our pilot subjects to iteratively compare a longer questionnaire derived from the TAM 2 survey with ours in

hopes of retaining acceptable content validity. This carries an implicit tradeoff in reduced construct validity.

The resulting survey tested for *Subjective Norm*, *Image*, *Job Relevance*, *Output Quality*, *Familiarity* and *Perceived Ease of Use* for the respondent. Variables measuring ChatGPT for faculty and student use respectively were *Perceived Usefulness* and *Result Demonstrability* (-F and -S variants), as well as a secondary *Perceived Ease of Use for Students* (under the assumption that PEOU extends to colleagues, but not necessarily to students). Figure 1 shows the structure of the survey, presented in full in appendix A.

The questionnaire was built using Qualtrics and contains an introduction and 4 blocks of questions. The question blocks were divided according to a logical flow rather than according to data requirements (Saunders et al., 2019).

Figure 2: Survey flow and Qualtrics block structure.



4.2.5. Variables

Below is a presentation of all measures included in our modelling and their associated items. All items were measured along a 7-point Likert scale (Strongly disagree to Strongly agree), as well as the reasoning. The quoted definitions of the measures are gathered from the TAM3 article by Venkatesh and Bala (2008) and includes their original reference.

Independent variables:

Subjective Norm (SN)

The degree to which an individual perceives that most people who are important to him think he should or should not use the system (Fishbein & Ajzen, 1975; Venkatesh & Davis, 2000).

A generic TAM item was selected based on to the interlinking of the higher educational context and the target populations' work environment:

SN: "My colleagues believe that using ChatGPT is a good idea."

Image (IMG)

The degree to which use of ChatGPT is perceived to enhance one's status in one's social system (Venkatesh & Davis, 2000).

A generic TAM item was selected based on the interlinking of the higher educational context and the target populations' professions:

IMG: "Using ChatGPT could enhance my professional image."

Job Relevance (JR)

The degree to which an individual believes that the target system is applicable to his or her job (Venkatesh & Davis, 2000).

A generic TAM item was selected based on the interlinking of the higher educational context and the target populations' work:

JR: "ChatGPT could be relevant to my work methods."

Output Quality (OR)

The degree to which an individual believes that the system performs his or her job tasks well (Venkatesh & Davis, 2000).

A generic TAM item was selected based on that the output of ChatGPT at the time of surveying was limited to information:

OR: "I perceive the information provided by ChatGPT to be reliable."

Perceived Ease of Use (PEOU)

The degree to which a person believes that using an IT will be free of effort (Davis, 1989).

Two generic TAM2 items were chosen pertaining to the general user-experience of ChatGPT as well as its learning curve, intended to be used as a construct:

PEOU: "I perceive ChatGPT to be user-friendly."

"I expect that learning to use ChatGPT will be straightforward."

Result Demonstrability for Faculty members/Students (RES-F and RES-S)

The degree to which an individual believes that the results of using a system are tangible, observable, and communicable (Moore & Benbasat, 1991).

A combination item was built by discussing "easily observable benefits" to capture the effects of tangibility within the concept of observability, with communicability assumed to be implicit. As we also deemed the likelihood of "benefits" associated with faculty versus student use of ChatGPT would carry different associations, the item was mirrored for the two groups:

RES-F: "The benefits of faculty using ChatGPT are easily observable."

RES-S: "The benefits of students using ChatGPT are easily observable."

Perceived Ease of Use for Students (PEOU-S)

Inspired by the thought that perceived Ease of Use for Students could be significantly different, we defined the following mirror question to one of the PEOU items to attempt to measure this effect:

PEOU-S: "I expect that learning to use ChatGPT will be straightforward."

Familiarity (FAM)

In an attempt to extrapolate the longitudinal effect of experience, we introduced the following item to measure familiarity with ChatGPT, derived from generic TAM formulation:

FAM: "I am familiar with ChatGPT or similar technologies."

Dependent variables:

Perceived Usefulness for Faculty members/Students (PU-F and PU-S)

The degree to which a person believes that using technology would enhance his or her job performance (Davis, 1989).

We recontextualized PU into two variables where we sought to capture the academic usefulness of ChatGPT in a qualitative sense and in an interaction-facilitating sense (Rowley, 1996). These items were then mirrored to form the respective measures:

PU-F: "I believe that using ChatGPT will enhance the quality of teaching."
"I believe that ChatGPT will facilitate faculty interaction with students."

PU-S: "I believe that using ChatGPT will enhance the quality of learning."
"I believe that ChatGPT will facilitate student interaction with faculty."

4.3. Research Method

4.3.1. Data Collection

The sample frame was generated through multi-stage sampling: The samples were naturally clustered by virtue of reaching out individually to all universities and university colleges (Swe. *Högskolor*) across Sweden asking for the correct point of contact for our request. Some participating schools elected to distribute it across all their institutions, whereas some asked us to contact their institutions individually. In total the survey was distributed across 24 institutions within seven universities and university colleges.

It should thus be noted that the sample is not truly random but has been pre-determined in scope by gatekeepers of the various institutions. The sampling method was chosen on the assumption that legitimacy through an intraorganizational mediator would be beneficial to response rate.

The data was collected between the 25th of October and 12th of November 2023. The Survey was opened 123 times, yielding 79 complete responses. Four of these responses were removed based on extreme answer patterns (max/min value answers only), significant deviation from average answer time, or non-researched roles yielding 75 valid survey responses which were used in our analysis. The participating institutions and included respondents fell under the following academic disciplines:

Table 1: Number of participating institutions and Valid responses by Academic Discipline

Academic Discipline	Number of Institutions	Valid responses
Social Sciences	9	30
Humanities	8	25
Natural sciences	4	14
Other	3	6

4.3.2. Quantitative Analysis Process

The sample data was collected from Qualtrics in an Excel spreadsheet upon the end of the data collection period. Analysis was conducted through statistical modelling with R through the integrated development environment RStudio.

The primary choice of model evaluation method is Ordinary Least Squares (OLS) regression. This is unlike the primary TAM studies, where analysis was conducted through Partial Least Squares Structural Equation Modelling (PLS-SEM). As our modelling features less complexity by virtue of primarily using single-item measures and not exploring complex hierarchical effects, OLS regression becomes applicable given that the underlying assumptions of the Gauss-Markov theorem holds.

We do however note that we are applying OLS regression on ordinal data. This is done primarily out of reasons of familiarity with the method and the scope of this thesis, albeit with admitted tradeoffs in ability to conclude from the model output (Winship & Mare, 1984). In order to combat this effect, a secondary non-parametric bootstrap model will be applied to test the main model assumptions.

4.4. Discussion of Method

4.4.1. Reliability, Validity, and Bias

Saunders et al. (2019) stresses the importance of the internal validity and reliability of the collected data. As described previously, we felt pressed to limit the number of questionnaire items to promote a higher response rate in a rather inaccessible demographic. Despite these efforts the overall number of responses, while sufficient for statistical analysis, are estimated to only represent a fraction of the studied population and are thought not to be truly random in nature. This has the following implications:

Reliability

Single-item measures carry an implicit risk of erroneous measurements being more pronounced in comparison to multi-item constructs, which mitigate these errors to a higher extent by averaging them out.

By having our pilot subjects evaluate the questions as our chosen validation method, we acknowledge a possible adverse effect on inter-rater reliability. While we saw a consensus among the pilot subjects, we acknowledge that this is a limited and subjective method and that none of the pilot subjects possessed expertise knowledge on TAM.

We still retain some constructs in our study, and thus have to assess their construct reliability. The retained constructs are constructed using two-item measures, which are inherently less reliable than multi-item scales and something to be mindful of in our analysis (Eisinger et al.,2012). Established literature posited that testing split-half reliability through the Spearman-Brown formula was the primary recommendation for two-item construct reliability measures (Eisinger et al., 2012; Saunders et al., 2019). Additionally, Cronbach Alpha is tested for completeness.

Validity

In terms of generalizability, it should be noted that efforts have been made to come in contact with nearly all academic institutions within Sweden. However, as our (implied) response rate is very low, we cannot guarantee that our sample is free of non-responder

bias and is a true average of the population. Likewise, as the participating institutions do not constitute a true random sample, it also has to be considered that other biases could be present in the data that risks rendering our conclusions inapplicable on the studied population (Saunders et al., 2019).

In terms of Content Validity, whether the measurement device provides adequate cover of the investigative questions, we ensured that we kept items for each of the constructs in TAM 2 up to Perceived Usefulness, in line with the established delimitation. The introduction of dual faculty and student measures is however a novel concept, and it is possible that they do not capture the effect we seek to evaluate.

An additional detriment to overall validity is that we did not to use attention checks. This was an unintentional omission due to an update to the Qualtrics form failing to apply. This was discovered after the survey had already been provided to several institutions, and the decision was taken to not alter the survey to ensure uniformity throughout distribution. While a lack of attention checks is associated with lower risk of Type II errors, the tradeoff is a significant rise in risk of Type I errors (Abbey & Meloy, 2017). We have attempted a rigorous analysis of survey completion times and multi-extreme value prevalence, omitting four responses as a consequence, but we cannot guarantee that this will not affect our outcomes or fully guarantee that the omissions truly represented inattentiveness.

Bias

The digital and anonymous nature of the survey, which respondents were informed of both in their initial email and upon commencing the survey, should minimize participant bias. The validated nature of TAM and the pure quantitative evaluation of the data should also minimize researcher bias, albeit one potential contributor is the process of selecting the single-item measures where subjective interpretations could have inadvertently affected the selection (Saunders et al., 2019).

5. Empirical Data

5.1. Descriptives

5.1.1. Sample demographics

See appendix B for descriptives of the sample demographic variables *Age*, *Gender*, *Academic Role*, and *Academic Discipline*. Among the 75 valid survey responses, age was distributed along a bell curve peaking in the 45-54 years range (30.7%). Gender skewed towards female (64.0%). 57.3% of respondents stated they were faculty members, followed by 28.0% researchers as the second most common role. Five out of six "Other" role respondents stated they were PhD students, and one stated their role to be head of department. Social science (40.0%) was the most commonly stated academic discipline, followed by Humanities (33.3%). Among the six "Other" discipline responses three respondents stated medicine, one stated education, one war science, and one a combination of natural and social science. 60 respondents, 80% of total, had previously used ChatGPT.

5.1.2. Construct Reliability

Table 2: Construct reliability tests, Spearman-Brown, Cronbach Alpha.

Construct	Spearman-Brown	Cronbach α	Items	Sample units
Perceived Usefulness Faculty (PU-F)	0.8652377	0.865	2	75
Perceived Usefulness Students (PU-S)	0.8849445	0.878	2	75
Perceived Ease of Use (PEOU)	0.6500182	0.648	2	75

Note: Accepted threshold for validity set at $p>0.7$.

Three constructs were established, PU-F, PU-S and PEOU. These were assessed for split-half reliability through Spearman-Brown and Cronbach Alpha as outlined in the methods section. PU-F and PU-S met the reliability criteria of $p\geq0.7$ for both tests. The intended PEOU construct failed to meet the criteria for split-half reliability at $p=0.65$. The constituent measures were retained as separate variables for the analysis to explore potential significance independently, with the measure pertaining to user-friendliness being redefined as User Friendliness (UF) and the measure on ease of learning (subjectively interpreted to be closer to the original TAM definition of perceived ease of use) retaining the label of PEOU.

5.1.3. OLS Regression model

The following model was the primary evaluation model:

$$(PU\text{–}F) = \beta_1(PEOU) + \beta_2(UF) + \beta_3(SN) + \beta_4(IMG) + \beta_5(JR) + \beta_6(QUAL) + \beta_7(RES\text{–}F) + \beta_8(FAM) + u_i$$

A mirror model of the primary evaluation model using the student measures was also evaluated:

$$(PU\text{–}S) = \beta_1(PEOU\text{–}S) + \beta_2(UF) + \beta_3(SN) + \beta_4(IMG) + \beta_5(JR) + \beta_6(QUAL) + \beta_7(RES\text{–}S) + \beta_8(FAM) + u_i$$

5.1.4. Dependent and Independent Variable descriptives

Table 3: DV and IV descriptives.

Variable	N	Mean	Median	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
1. Perceived Usefulness Faculty (PU-F)	75	3.407	3.5	1.447	1	2	4.5	7
2. Perceived Ease of Use (PEOU)	75	4.88	5	1.395	1	4	6	7
3. Result Demonstrability Faculty (RES-F)	75	3.16	3	1.611	1	2	4	7
4. Perceived Usefulness Students (PU-S)	75	3.107	3	1.413	1	2	4	7
5. Perceived Ease of Use Students (PEOU-S)	75	4.387	5	1.739	1	3	6	7
6. Result Demonstrability Students (RES-S)	75	2.933	3	1.446	1	2	4	7
7. User friendliness (UF)	75	5.093	5	1.265	1	4	6	7
8. Familiarity (FAM)	75	5.093	6	1.403	1	5	6	7
9. Subjective Norm (SN)	75	3.654	4	1.333	1	3	4	7
10. Image (IMG)	75	3.974	4	1.672	1	3	5	7
11. Output Quality (QUAL)	75	2.987	3	1.428	1	2	4	7
12. Job Relevance (JR)	75	4.538	5	1.639	1	4	5.5	7

Variables were checked for skewness and kurtosis (table C1, appendix C). Four items deviated from a normal distribution. A Shapiro-Wilks test on the dependent variables PU-F, PEOU, PU-S, and PEOU-S, showed that these too significantly deviate from a normal distribution. This does not have major implications on the regression analysis given that the residuals are normally distributed. Variables are visualized in a boxplot in figure C1, appendix C.

Table 4: Correlation table, DVs and IVs.

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. Perceived Usefulness Faculty (PU-F)	1											
2. Perceived Ease of Use (PEOU)	0.309 **	1										
3. Result Demonstrability Faculty (RES-F)	0.786 ***	0.370 **	1									
4. Perceived Usefulness Students (PU-S)	0.863 ***	0.329 **	0.738 ***	1								
5. Perceived Ease of Use Students (PEOU-S)	0.297 **	0.398 ***	0.436 ***	0.308 **	1							
6. Result Demonstrability Students (RES-S)	0.705 ***	0.204	0.770 ***	0.725 ***	0.419 ***	1						
7. User Friendliness (UF)	0.304 **	0.482 ***	0.331 **	0.395 ***	0.303 **	0.225	1					
8. Familiarity (FAM)	0.294 *	0.464 ***	0.353 **	0.365 **	0.128	0.235 *	0.536 ***	1				
9. Subjective Norm (SN)	0.234 *	0.172	0.286 *	0.291 *	0.151	0.344 **	0.133	0.056	1			
10. Image (IMG)	0.563 ***	0.183	0.404 ***	0.620 ***	0.029	0.368 **	0.309 **	0.217	0.430 ***	1		
11. Output Quality (QUAL)	0.595 ***	0.115	0.430 ***	0.580 ***	0.078	0.458 ***	0.248 *	0.143	0.082	0.435 ***	1	
12. Job Relevance (JR)	0.541 ***	0.328 **	0.533 ***	0.564 ***	0.198	0.516 ***	0.395 ***	0.371 **	0.358 **	0.604 ***	0.540 ***	1

Note: * ρ <0.05; ** ρ <0.01; *** ρ < 0.001. N=75

All independent variables are independently significantly correlated with the for analysis primary dependent variables of PU-F and PU-S. Variance Inflation Factor measurements were established for each regression, with no value exceeding five throughout the analysis, ensuring lack of multicollinearity. These are attached to the appendix entries for each model.

PU-F and PU-S show the highest correlation within the sample. Ahead of regression modelling we elected to see if there was a significant difference between the Faculty and Student measurements. As they were previously proven to be non-normal, this was done through a non-parametric Wilcoxon signed-rank test.

Table 5: Non-parametric Wilcoxon signed-rank test between Faculty/Student mirror variables, full sample.

Variable	V:	P-value:	95% CI Lower	95% CI Upper	(Pseudo)Median:
1. Perceived Usefulness Faculty vs Student	811.5	0.000757	0.249956	0.7500709	0.5000403
2. Perceived Ease of Use Personal vs Student	878	0.01848	3.09433e-06	1.499952e+00	0.9999322
3. Result Demonstrability Faculty vs Student	494	0.06032	- 3.026083e-05	1.000039e+00	0.5

Result demonstrability could not be determined to be significantly different between the faculty and student measure at p = 0.05, albeit this does not affect the research model. Most notable is the finding that PU-F and PU-S are significantly different.

5.2. Regression Modelling Outcomes

5.2.1. Perceived Usefulness for Faculty

Table 6: OLS Model, Perceived Usefulness (PU-F), full sample (N=75)

PU-F	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.858e-16	8.984e-02	0.000	1.00000
PEOU	5.480e-02	7.974e-02	0.687	0.49434
UF	-4.624e-02	9.212e-02	-0.502	0.61737
SN	-6.210e-02	7.906e-02	-0.786	0.43496
IMG	2.358e-01	7.315e-02	3.224	0.00197 **
JR	-6.504e-02	8.354e-02	-0.779	0.43906
QUAL	2.640e-01	8.081e-02	3.267	0.00173 **
RES-F	5.481e-01	7.188e-02	7.626	1.24e-10 ***
FAM	9.809e-03	8.240e-02	0.119	0.90561

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.778 on 66 degrees of freedom
Multiple R-squared: 0.742, Adjusted R-squared: 0.7107
F-statistic: 23.73 on 8 and 66 DF, p-value: < 2.2e-16

Note: A full breakdown of the model, VIF, and residual analysis can be found in appendix D.

Initial modelling indicated that three variables had a significant (and positive) relationship with Perceived Usefulness for Faculty: Image, Output Quality, and Result Demonstrability for Faculty. Most notable was the lack of correlation for Perceived Ease of Use (and the closely related extension User Friendliness), which contradicts the underlying assumptions of TAM. Despite this the model maintains a relatively high adjusted R^2 value at 0.7107. VIF indicates no multicollinearity. Shapiro-Wilk and Breusch-Pagan tests indicates normality and heteroscedasticity assumptions holds on residuals.

A test of PEOU despite its non-significance can be found in appendix E, where it was found that User-Friendliness was the only significant contributor. Further testing on PU-S as well as subsamples maintained that PEOU or PEOU-S held no significance for the sample population. This means a full rejection of hypothesis 1:

Perceived Ease of Use

H1	Perceived Ease of Use has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use.	**Not supported**

5.2.2. Mediation through Image

Modelling the other variables on IMG gave the following result:

Table 7: OLS Model, Mediation of other variables through Image (IMG), full sample (N=75)

Image	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4.714e-17	1.500e-01	0.000	1.00000
PEOU	-8.501e-02	1.328e-01	-0.640	0.52420
UF	1.328e-01	1.530e-01	0.868	0.38865
SN	3.359e-01	1.255e-01	2.677	0.00934 **
JR	3.792e-01	1.316e-01	2.881	0.00532 **
QUAL	2.061e-01	1.326e-01	1.554	0.12479
RES-F	4.737e-02	1.199e-01	0.395	0.69403
FAM	2.654e-03	1.376e-01	0.019	0.98467

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.299 on 67 degrees of freedom
Multiple R-squared: 0.4532, Adjusted R-squared: 0.396
F-statistic: 7.932 on 7 and 67 DF, p-value: 5.536e-07

Note: A full breakdown of the model, VIF, and residual analysis can be found in appendix F.

Modelling the mediation effect through showed significant (and positive) mediation through Image not only of Subjective Norm, as predicted by the TAM framework, but also of Job Relevance. These mediation effects were confirmed to be significant through Sobel testing (appendix G).

While testing indicates the OLS assumptions held for all our models, we elected to model both the primary and mediating models through bootstrapping (10,000 resamples, appendix H) for further effect confirmation. The bootstrapped model confirmed the same mediation effects, with the addition of Familiarity being significant in the Primary model and Output Quality instead being mediated through Image rather than directly significant.

5.2.3. Perceived Usefulness for Students

Table 8: OLS Regression results, PU-S.

PU-S	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-6.121e-17	8.916e-02	0.000	1.0000
PEOU-S	5.393e-02	6.098e-02	0.885	0.3796
UF	8.147e-02	9.198e-02	0.886	0.3790
SN	-4.902e-02	7.990e-02	-0.614	0.5416
IMG	3.203e-01	7.356e-02	4.354	4.76e-05 ***
JR	-8.873e-02	8.293e-02	-1.070	0.2886
QUAL	2.090e-01	8.214e-02	2.544	0.0133 *
RES-S	4.712e-01	8.516e-02	5.534	5.81e-07 ***
FAM	1.337e-01	7.930e-02	1.686	0.0965 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7722 on 66 degrees of freedom
Multiple R-squared: 0.7335, Adjusted R-squared: 0.7012
F-statistic: 22.7 on 8 and 66 DF, p-value: 3.258e-16

Note: A full breakdown of the model, VIF, and residual analysis can be found in appendix I.

The mirror model on PU-S showed a similar non-significance for PEOU-S as PEOU on PU-F, and by nature of sharing many underlying variables the mediation effect of SN and JR through IMG. It too holds a relatively high adjusted R^2 of 0.7012.

5.2.4. Subsample testing

The somewhat restrictive number of respondents meant options for effective subsampling was limited. The primary model for PU-F and the PU-S model was non-the-less tested on two subsamples, one drawn only from previous users of ChatGPT (N=60, appendix J, K & M), and one drawn from the faculty member respondents (N=43, appendix N, O & P),

Attempts were made at subsampling the intersection of only faculty members that had previously used ChatGPT (N=36), but the normality assumptions for residuals were rejected and it is not included in this paper. Likewise, due to the limited sample size the opposite sample group of respondents without ChatGPT experience could not be tested.

The significant and relationship positive influence of IMG, QUAL and RES-F on PU-F was maintained on all subsamples. Specifically in the Faculty only subsample, the PU-S model saw that QUAL on PU-S fell out of significance with P-values in the 0.05 to 0.10 range.

The mediation effect through Image had one notable deviation per subsample: the Breusch-Pagan test failed to reject the homoscedasticity null-hypothesis in the previous usage sample but was otherwise congruent with the full sample finding. In the faculty only sample all conditions were met, but SN was found to be non-significant.

6. Analysis

6.1. Descriptives

With a 7-point Likert scale, a value of 4 represents a sentiment on the average of the scale, and the maximum achievable standard deviation is 3. As the scale is ordinal in nature, and several variables were found to be non-normally distributed, we avoid making inferences based on the numerical values beyond observing that the observations see a widespread in answer range as both the standard deviations and boxplot (figure C1) shows.

6.2. Model evaluation

6.2.1. Variance explanation

The primary faculty model and supplementary student model achieve adjusted R^2 values of 0.7107 and 0.7012 respectively, indicating that they explain 71% and 70% of the variance present in PU-F and PU-S. This indicates a good fit in between both model iterations, albeit we withhold comparison with external data due to the ordinal scale.

6.2.2. Effect size

Consistent throughout the modelling is that Result Demonstrability has the largest effect (0.548) on Perceived Usefulness in Faculty use, with an effect size over twice that of Image (0.264) and Output Quality (0.236). Notable is that when delimited to only Faculty members, the effect sizes were much closer in value overall (0.433, 0.290, 0.360). In regard to the mediating effect through Image, both Subjective Norm (0.336) and Job Relevance (0.379) had similar effect size in the overall sample, but Job Relevance is the major (and only significant) determinant in the Faculty group (0.486). Perceived Usefulness in Student use saw similar effect size patterns. We reiterate that this analysis is made on ordinal data, and that conclusions drawn from this should be made cautiously, albeit effect size appeared similar in the non-parametric bootstrap.

6.2.3. Validation and model assumptions

Visual inspection of residual plots does not give rise to concern, especially for the full sample tests. In all models tested, assumptions of normality of residuals through Spearman-Wilk test held true. Breusch-Pagan testing indicated a potential issue on one subsample model of image mediation, but no other model presented evidence of heteroskedasticity. Furthermore, variance inflation factors indicate no notable signs of multicollinearity with the highest observation across models being 2.41. These factors validate the regression models.

6.3. Hypothesis evaluation

An initial conclusion in the findings is that there is no support for hypothesis 1, i.e. there is no significant relationship between Perceived Ease of Use and Perceived Usefulness

(for either of the faculty and student models). While this does not mean a complete rejection of Perceived Ease of Use as a determinant for Behavioral Intent, which is not observed in this study, it does imply a deviation from TAM's normal assumption. It should also be stated that the failure to achieve a viable construct and further deconstruction into two items could have exaggerated this outcome.
In terms of effects that can be confirmed to be directly positive and significant on Perceived Usefulness, we observe three findings: Image, Output Quality and Result Demonstrability:

H4	Image has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use.	**Supported**
H6	Output Quality has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use.	**Partially supported**
H7	Result Demonstrability has a direct, positive. and significant correlation with Perceived Usefulness of ChatGPT in faculty use.	**Supported**

The above hypotheses are affirmed in all tested scenarios, except for that the supplementary bootstrap model showed an indirect (via Image) rather than direct effect for Output Quality. It is therefore noted as partially supported.

Additionally, the variables Subjective Norm and Job Relevance were found to not have direct effects. The indirect effect of Subjective Norm mediated through Image was however present, and furthermore a similar effect could be identified for Job Relevance:

H2	Subjective Norm has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use	**Not supported**
H3	Subjective Norm has an indirect, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use through mediation via Image.	**Supported**
H5	Job Relevance has a direct, positive, and significant correlation with Perceived Usefulness of ChatGPT in faculty use.	**Not supported**

Notable is that Subjective Norm failed to see a mediated effect via Image when only Faculty respondents were included in the modelling.

Both hypotheses informing the model extension concerning faculty use versus student use are supported by the data. The Wilcoxon signed-rank test affirmed a significant difference between perceived usefulness for faculty versus students, whereas the models affirmed that the same independent variables are significant when measuring against either dependent variable:

H8	Perceived Usefulness of ChatGPT in student use is significantly different within the sampled population compared to Perceived Usefulness of ChatGPT in faculty use.	**Supported**
H9	Perceived Usefulness of ChatGPT in student use shares the underlying significant independent variables that present for Perceived Usefulness of ChatGPT in faculty use.	**Supported**

6.4. Analysis of sample demographics

Comparing to demographics among the population of higher education employees in Sweden (SCB, 2021), the sample aligns closely in age distribution (77% aged 35-64, median 45-49) and fairly close along the Faculty/Researcher divide, but less so in gender (54% male) and academic discipline where Medicine is underrepresented and Humanities overrepresented. This indicates that the sample is not fully representative of the overall population of academic faculty and staff in Sweden, which could have implications for the generalizability of our findings when looking at the full population.

7. Discussion

7.1. Answering the Research Question

This thesis set out to answer the research question:

What underlying factors hold significant explanatory value for acceptance of Artificial Intelligence in Academia among academic faculty and staff, in accordance with the Technology Acceptance Model 2?

Under the assumption that the, via the body of research surrounding TAM, established valid connection between *Perceived Usefulness* and acceptance of information systems holds true, the following answer can be provided:

Informed through the Technology Acceptance Model 2, our data supports that *Image* and *Result Demonstrability* hold direct positive and significant explanatory value for the established determinant of acceptance *perceived usefulness*. The data also support a significant positive but indirect explanatory value from *Subjective Norm* and *Job Relevance* mediated through *Image*, with *Output Quality* falling into either category dependent on regression method. The data fails to support that *Perceived Ease of Use* has significant explanatory value for *Perceived Usefulness*, but as the study has not directly examined its effect on acceptance such as originally posited by the TAM framework, a conclusive statement cannot be made.

7.2. Study contributions

Theoretical implications

Acceptance of ChatGPT in Academia is a young research topic. At the point of work commencing there was no prior research utilizing the Technology Acceptance Model framework to map perceived usefulness of ChatGPT among academic faculty and staff. Aimed at adding contrast to the heavily researched student body (e.g., Tiwari et al., 2023; Zou et al., 2023), the thesis affirms that most of the underlying assumptions of TAM2 on perceived usefulness are applicable on the population of academic faculty and staff when researching the use of ChatGPT, albeit in a novel way where some determinants are mediated and perceived ease of use being a non-significant factor.

The finding that there is a significant difference in views of usefulness of ChatGPT when used by students versus faculty is also a confirmed phenomenon without prior evidence. Additionally, that the faculty group has notable effect size differences from the full sample suggests that there exist discrepancies between academic faculty and other staff in their views of ChatGPT usage within academia, and that future research could be serviced by stratification of the groups.

Practical implications

The findings of this study can contribute to the ongoing discourse surrounding ChatGPT in academia and informed decision making in the academic sphere pertaining to the implementation of policies concerning the use of ChatGPT and generative AI within academic institutions.

In terms of the faculty-student mirror variables, our data shows that the questions pertaining to perceived usefulness and result demonstrability all saw answers towards the disagreement side of the scale. This indicates a slight negative skew, which contrasts with the findings of the literature review where sentiments were overall positive (i.e. Bringula, 2023). This could have explanatory value in the slow adoption of artificial intelligence as identified by Kasneci et al. (2023). Potential alternative explanations of this finding could be an effect endemic to the thesis's geographical delimitation of Sweden, or pertaining to other demographic factors present within the sample, albeit this would require further study.

The high significance of result demonstrability's effect on perceived usefulness could potentially implicate that acceptance of ChatGPT among academic faculty and staff will increase as the capabilities of artificial intelligence continues to develop. If this were the case one could infer a higher likelihood of voluntary adoption of ChatGPT in academic settings in the future.

7.3. Research Design and Limitations

As has been thoroughly discussed in this thesis, several tradeoffs were made in hopes of expediency generating a larger set of respondents from the surveying process, and OLS regression was applied on the ordinal dataset. This has implications for the validity of the findings, and we recommend caution when inferring from our data and findings and suggest that it is used to inform for but not deduce findings of future research. We also advise against its inclusion in meta-analyses.

Other limitations include the restricted sample size of 75 observations, where the assumed population ranges in the thousands. The geographic delimitation to the region of Sweden also means that the findings could be less generalizable when applied in an international context, especially outside of northern Europe.

A final limitation is that this thesis is built upon data and literature primarily collected in the period of October and November of 2023, but is published in May 2024. This implies that new additions could have been made to the body of research in the meantime that more accurately present the current conditions of the research area. The largest discrepancy and source of uncertainty is the release of the updated ChatGPT-4 model shortly after the end of our data collection in November 2023. This new edition saw a multitude of improvements, and with its improved capabilities can be assumed to have shifted sentiment favorably.

7.4. Recommendations for Future Research

This thesis has found initial evidence of significant factors underlying the acceptance of ChatGPT among academic faculty and staff. The (since previously) validated nature of the Technology Acceptance Model framework makes this an incremental addition to the body of research and serves at most as specific and explicit confirmation of a previously inferred phenomenon. The nature of this research as discussed above implicates that a more in-depth study applying the Technological Acceptance Model framework could be advisable. This should include a survey and modelling with fully sized construct measurements, leveraging partial least squares (PLS) structural equation modelling for thorough hierarchical analysis. A longitudinal study to observe effects over time would also be advisable.

Some suggestions for future research vectors outside the scope of testing TAM's applicability, based on our findings, would be:

- Qualitative study of how Image, Output Quality and/or Result Demonstrability drive acceptance of ChatGPT in academia.
- Investigating alternative reasons behind the slow introduction of ChatGPT in academia.

- A deeper study of the discrepancy between views on students versus faculty use of ChatGPT.

8. References

Ajzen, I., & Fishbein, M. (2000). Attitudes and the Attitude-Behavior Relation: Reasoned and Automatic Processes. European Review of Social Psychology, 11(1), 1–33. https://doi.org/10.1080/14792779943000116

Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single Item Measures in Psychological Science. European Journal of Psychological Assessment. https://doi.org/10.1027/1015-5759/a000699

Bahroun, Z., Tanash, M., As'ad, R., & Alnajar, M. (2023). Artificial intelligence applications in project scheduling: A systematic review, bibliometric analysis, and prospects for future research. Management Systems in Production Engineering, 31(2), 144-161. https://doi.org/10.2478/mspe-2023-0017

Bação, P., Lopes, V. G., & Simões, M. (2023). AI, demand and the impact of productivity-enhancing technology on jobs: Evidence from Portugal. Eastern European Economics, 61(4), 353–377. https://doi.org/10.1080/00128775.2022.2064307

Bringula, R. (2023). What do academics have to say about ChatGPT? A text mining analytics on the discussions regarding ChatGPT on research writing. AI Ethics. https://doi.org/10.1007/s43681-023-00354-w

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. International Journal of Educational Technology in Higher Education, 20, 1-22. https://doi.org/10.1186/s41239-023-00392-8

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

Eisinga, R., Grotenhuis, M., & Pelzer, B. (2012). The reliability of a two-item scale: Pearson, Cronbach or Spearman-Brown? International Journal of Public Health. http://dx.doi.org/10.1007/s00038-012-0416-3

Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2021). Artificial intelligence and business value: A literature review. Information Systems Frontiers, 24, 1709-1734. https://doi.org/10.1007/s10796-021-10186-w

Hasanein, A. M., & Abu. (2023). Drivers and Consequences of ChatGPT Use in Higher Education: Key Stakeholder Perspectives. European Journal of Investigation in Health, Psychology and Education, 13(11), 2599–2614. https://doi.org/10.3390/ejihpe13110181

Higher Education. Employees in Higher Education 2020. Sveriges officiella statistic Statistiska Meddelanden UF 23 SM 2101. https://www.scb.se/contentassets/e3e2c2f3c5574ef3847db70c04bcbacf/uf0202_2020a01_sm_uf23sm2101.pdf

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnermann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdle, C., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT impact academia and libraries? Library Hi Tech News, 40(3), 26-29. https://doi.org/10.1108/LHTN-01-2023-0009

Meyer, J. G., Urbanowicz, R. J., Martin, P. C. N., & others. (2023). ChatGPT and large language models in academia: opportunities and challenges. BioData Mining, 16, 20. https://doi.org/10.1186/s13040-023-00339-9

Mills, A., Bali, M., & Eaton, L. (2023). How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process. Journal of Applied Learning & Teaching, 6(1), 1-17. https://doi.org/10.37074/jalt.2023.6.1.34

Rossiter, J. R., & Braithwaite, B. (2013). C-OAR-SE-Based Single-Item Measures for the Two-Stage Technology Acceptance Model. Australasian Marketing Journal. https://doi.org/10.1016/j.ausmj.2012.08.005

Saunders, M. N. K., Lewis, P., & Thornhill, A. (n.d.). Research Methods for Business Students (8th edition). ISBN: 9781292208787.

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. Journal of Applied Learning & Teaching, 6(1), 1-17. https://doi.org/10.37074/jalt.2023.6.1.17

Tiwari, C. K., Bhat, M. A., Khan, S. T., Subramaniam, R., & Khan, M. A. I. (2023). What drives students toward ChatGPT? An investigation of the factors influencing adoption and usage of ChatGPT. Interactive Technology and Smart Education. https://doi.org/10.1108/ITSE-04-2023-0061

Venkatesh, V., Brown, S. A., Maruping, L. M., & Bala, H. (2008). Predicting Different Conceptualizations of System Use: The Competing Roles of Behavioral Intention, Facilitating Conditions, and Behavioral Expectation. MIS Quarterly, 32(3), 483–502. https://doi.org/10.2307/25148853

Weber, M., Engert, M., Schaffer, N., Weking, J., & Krcmar, H. (2022). Organizational Capabilities for AI Implementation—Coping with Inscrutability and Data Dependency in AI. Information Systems Frontiers. https://doi.org/10.1007/s10796-022-10297-y

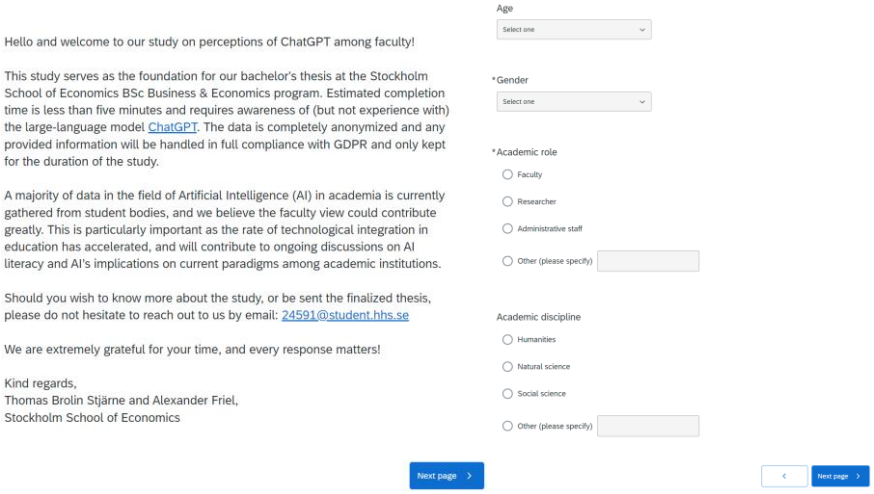
Winship, C., & Mare, R. D. (1984). Regression Models with Ordinal Variables. American Sociological Review, 49(4), 512–525. https://doi.org/10.2307/2095465

Yousafzai, S. Y., Foxall, G. R., & Pallister, J. G. (2007). Technology acceptance: A meta-analysis of the TAM: Part 1. Journal of Modelling in Management, 2(3). https://doi.org/10.1108/17465660710834453

Zou, M., & Huang, L. (2023). To use or not to use? Understanding doctoral students' acceptance of ChatGPT in writing through technology acceptance model. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1259531

9. Appendices

Appendix A: Full survey



Hello and welcome to our study on perceptions of ChatGPT among faculty!

This study serves as the foundation for our bachelor's thesis at the Stockholm School of Economics BSc Business & Economics program. Estimated completion time is less than five minutes and requires awareness of (but not experience with) the large-language model ChatGPT. The data is completely anonymized and any provided information will be handled in full compliance with GDPR and only kept for the duration of the study.

A majority of data in the field of Artificial Intelligence (AI) in academia is currently gathered from student bodies, and we believe the faculty view could contribute greatly. This is particularly important as the rate of technological integration in education has accelerated, and will contribute to ongoing discussions on AI literacy and AI's implications on current paradigms among academic institutions.

Should you wish to know more about the study, or be sent the finalized thesis, please do not hesitate to reach out to us by email: 24591@student.hhs.se

We are extremely grateful for your time, and every response matters!

Kind regards,
Thomas Brolin Stjärne and Alexander Friel,
Stockholm School of Economics

Age
Select one

*Gender
Select one

*Academic role
- Faculty
- Researcher
- Administrative staff
- Other (please specify)

Academic discipline
- Humanities
- Natural science
- Social science
- Other (please specify)

Next page

*I have previously used ChatGPT (any amount of private or professional use)
- Yes
- No

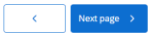
*My workplace has a central policy in place regarding ChatGPT use (concerning student use, faculty use, or both)
- Yes
- No
- Do not know

*Personal perceptions of ChatGPT

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I am familiar with ChatGPT or similar technologies.							
I feel that using ChatGPT is entirely my choice.							
I perceive ChatGPT to be user-friendly.							
I expect that learning to use ChatGPT will be straightforward.							
My colleagues believe that using ChatGPT is a good idea.							
Using ChatGPT could enhance my professional image.							
I perceive the information provided by ChatGPT to be reliable.							
ChatGPT could be relevant to my work methods.							

*Perceptions of faculty's usage of ChatGPT

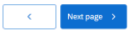
	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I believe that using ChatGPT will enhance the quality of teaching.							
I believe that ChatGPT will facilitate faculty interaction with students.							
The benefits of faculty using ChatGPT are easily observable.							



Previous page | Next page

*Perceptions of students' usage of ChatGPT

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I believe that using ChatGPT will enhance the quality of students' learnings.							
I believe that ChatGPT will facilitate student interaction with faculty.							
The benefits of students using ChatGPT are easily observable.							
I expect that learning to use ChatGPT will be straightforward for students.							



Previous page | Next page

Appendix B: Demographics

Table B1: Demographics breakdown, N=75

Variable	N	% of total
Age		
Under 25:	0	0.00%
25-34:	11	14.7%
35-44:	16	21.3%
45-54:	23	30.7%
55-64:	15	20.0%
Over 64:	10	13.3%
Gender		
Female:	48	64.0%
Male:	24	32.0%
Non-binary:	0	0.00%
Prefer not to say:	3	4.0%
Academic role		
Faculty:	43	57.3%
Researcher:	21	28.0%
Administrative staff:	5	6.7%
Other:	6	8%
Academic discipline		
Humanities:	25	33.3%
Social Science:	30	40.0%
Natural Science:	14	18.7%
Other:	6	8.0%
Has used ChatGPT		
Yes:	60	80%
No:	15	20%

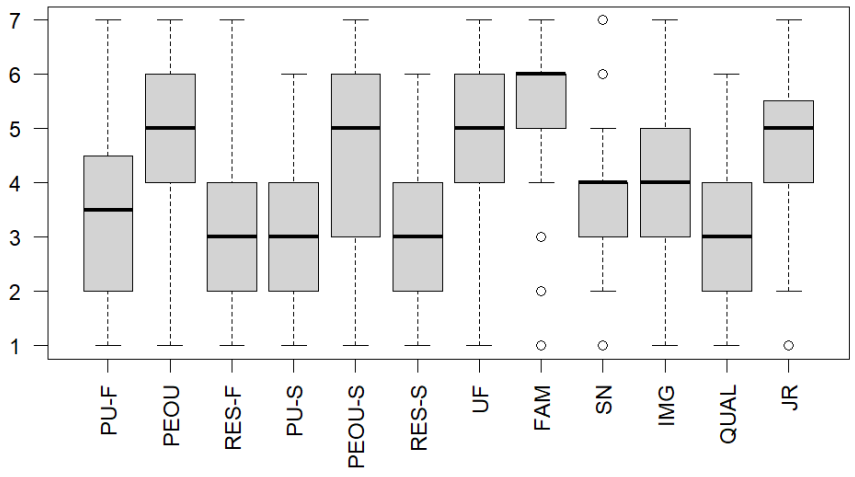
Appendix C: Variable analysis

Table C1: Skewness and Kurtosis testing of variables.

Variable	Skewness	Kurtosis	Jarque-Bera	JB p-value
1. Perceived Usefulness Faculty (PU-F)	0.12738035	-0.8111306	1.9763	0.3723
2. Perceived Ease of Use (PEOU)	-0.52629420	-0.2182014	3.6681	0.1598
3. Result Demonstrability Faculty (RES-F)	0.60417044	-0.7060425	6.0448	0.04868 *
4. Perceived Usefulness Students (PU-S)	0.02678700	-1.2144173	4.2567	0.119
5. Perceived Ease of Use Students (PEOU-S)	-0.42828963	-1.0495247	5.49	0.06425
6. Result Demonstrability Students (RES-S)	0.22043219	-0.9605535	3.1921	0.2027
7. User friendliness (UF)	-0.44959782	0.1604994	2.8204	0.2441
8. Familiarity (FAM)	-1.17745745	0.7897956	20.534	3.477e-05 **
9. Voluntariness (VOL)	-0.59330055	-0.8702533	6.6429	0.0361 *
10. Subjective Norm (SN)	0.11945811	-0.3576691	0.44091	0.8022
11. Image (IMG)	-0.09176017	-0.7142605	1.4383	0.4872
12. Output Quality (QUAL)	0.35244608	-0.8126794	3.3892	0.1837
13. Job Relevance (JR)	-0.71535986	-0.4411528	7.0909	0.02885 *

Note: JB test rejects null hypothesis of skewness and kurtosis matching normal distribution if $p<0.05$.

Figure C1: Box plot visualization of variables



Appendix D: Modelling Perceived Usefulness for Faculty (PU-F), full sample (N=75)

Table D1: OLS Model, Perceived Usefulness (PU-F), full sample (N=75)

PU-F	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.858e-16	8.984e-02	0.000	1.00000
PEOU	5.480e-02	7.974e-02	0.687	0.49434
UF	-4.624e-02	9.212e-02	-0.502	0.61737
SN	-6.210e-02	7.906e-02	-0.786	0.43496
IMG	2.358e-01	7.315e-02	3.224	0.00197 **
JR	-6.504e-02	8.354e-02	-0.779	0.43906
QUAL	2.640e-01	8.081e-02	3.267	0.00173 **
RES-F	5.481e-01	7.188e-02	7.626	1.24e-10 ***
FAM	9.809e-03	8.240e-02	0.119	0.90561

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.778 on 66 degrees of freedom
Multiple R-squared: 0.742, Adjusted R-squared: 0.7107
F-statistic: 23.73 on 8 and 66 DF, p-value: < 2.2e-16

Table D2: Variance Inflation Factors, Perceived Usefulness for Faculty (PU-F), full sample

Variable	VIF
PEOU	1.511936
UF	1.658997
SN	1.358277
IMG	1.828697
JR	2.291452
QUAL	1.628899
RES-F	1.639374
FAM	1.634050

Residual analysis, Perceived Usefulness for Faculty (PU-F), full sample

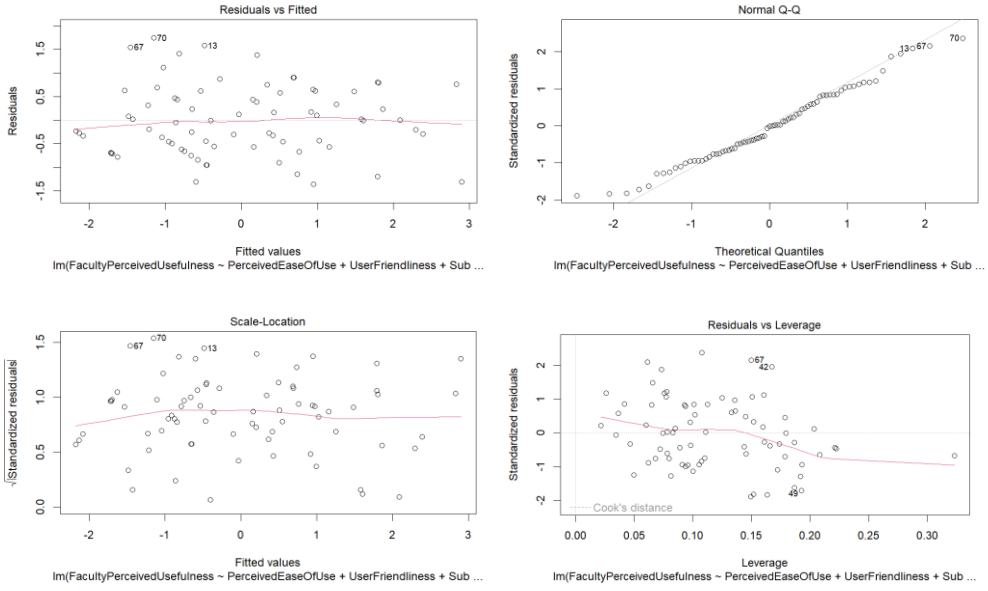
Table D3: Breusch-Pagan test, null-hypothesis of homoscedastic residuals cannot be rejected if p<0.05.

BP	df	p-value
8.3801	8	0.3972

Table D4: Shapiro-Wilk normality test, null-hypothesis of normally distributed residuals cannot be rejected if p<0.05.

W	p-value
0.97961	0.268

Figure D1: Residuals visualizations, OLS Model, Perceived Usefulness (PU-F), full sample



Appendix E: Modelling Perceived Ease of Use (PEOU), full sample (N=75)

Table E1: OLS Model, Perceived Ease of Use (PEOU), full sample (N=75)

PU-F	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.559e-16	1.376e-01	0.000	1.0000
UF	3.270e-01	1.354e-01	2.416	0.0184 *
SN	7.624e-02	1.208e-01	0.631	0.5300
IMG	-7.153e-02	1.117e-01	-0.640	0.5242
JR	8.729e-02	1.275e-01	0.684	0.4961
QUAL	-9.58e-02	1.232e-01	-0.777	0.4397
RES-F	1.664e-01	1.082e-01	1.538	0.1288
FAM	2.263e-01	1.232e-01	1.837	0.0707 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.192 on 67 degrees of freedom
Multiple R-squared: 0.3386, Adjusted R-squared: 0.2695
F-statistic: 4.9 on 7 and 67 DF, p-value: 0.0001652

Table E2: Variance Inflation Factors, Perceived Usefulness for Faculty (PU-F), full sample

Variable	VIF
UF	1.526042
SN	1.350244
IMG	1.817577
JR	2.275542
QUAL	1.614340
RES-F	1.583489
FAM	1.555709

Residual analysis, Perceived Usefulness for Faculty (PU-F), full sample

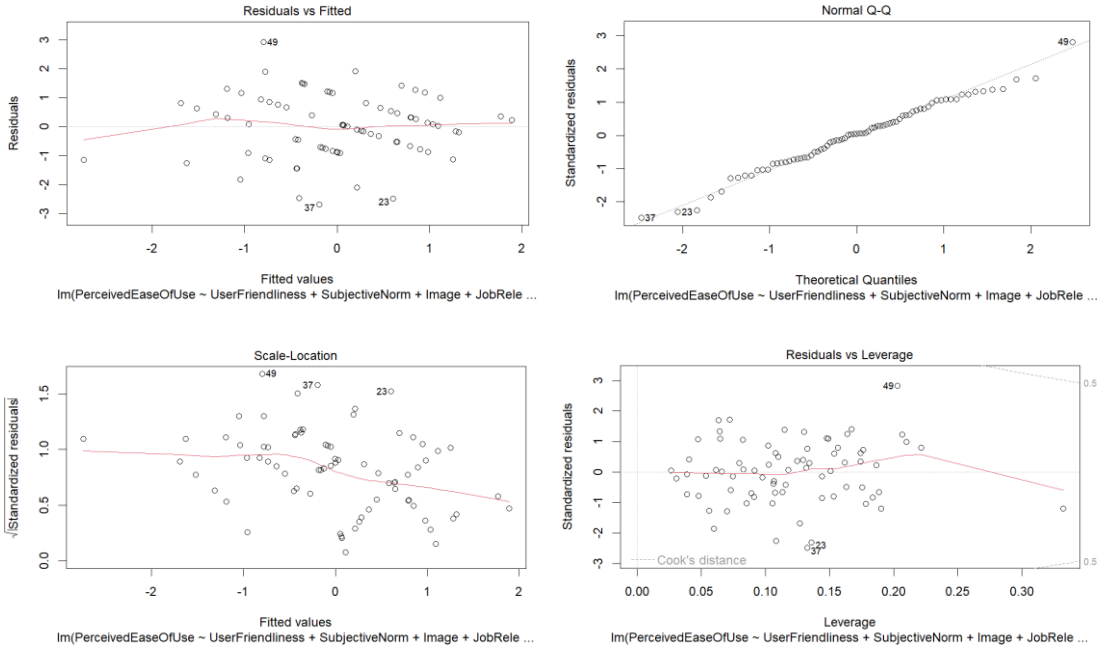
Table E3: Breusch-Pagan test, null-hypothesis of homoscedastic residuals cannot be rejected if p<0.05.

BP	df	p-value
10.582	7	0.1579

Table E4: Shapiro-Wilk normality test, null-hypothesis of normally distributed residuals cannot be rejected if p<0.05.

W	p-value
0.97086	0.08101

Figure E1: Residuals visualizations, OLS Model, Perceived Ease of Use (PEOU), full sample



Appendix F: Modelling Mediation of other variables through Image (IMG) on PU-F, full sample (N=75)

Table F1: OLS Model, Mediation of other variables through Image (IMG), full sample (N=75)

Image	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4.714e-17	1.500e-01	0.000	1.00000
PEOU	-8.501e-02	1.328e-01	-0.640	0.52420
UF	1.328e-01	1.530e-01	0.868	0.38865
SN	3.359e-01	1.255e-01	2.677	0.00934 **
JR	3.792e-01	1.316e-01	2.881	0.00532 **
QUAL	2.061e-01	1.326e-01	1.554	0.12479
RES-F	4.737e-02	1.199e-01	0.395	0.69403
FAM	2.654e-03	1.376e-01	0.019	0.98467

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.299 on 67 degrees of freedom
Multiple R-squared: 0.4532, Adjusted R-squared: 0.396
F-statistic: 7.932 on 7 and 67 DF, p-value: 5.536e-07

Table F2: Variance Inflation Factors, Mediation of other variables through Image (IMG)

Variable	VIF
PEOU	1.502742
UF	1.640561
SN	1.227062
JR	2.038830
QUAL	1.572199
RES-F	1.635564
FAM	1.634041

Residual analysis, Mediation of other variables through Image (IMG)

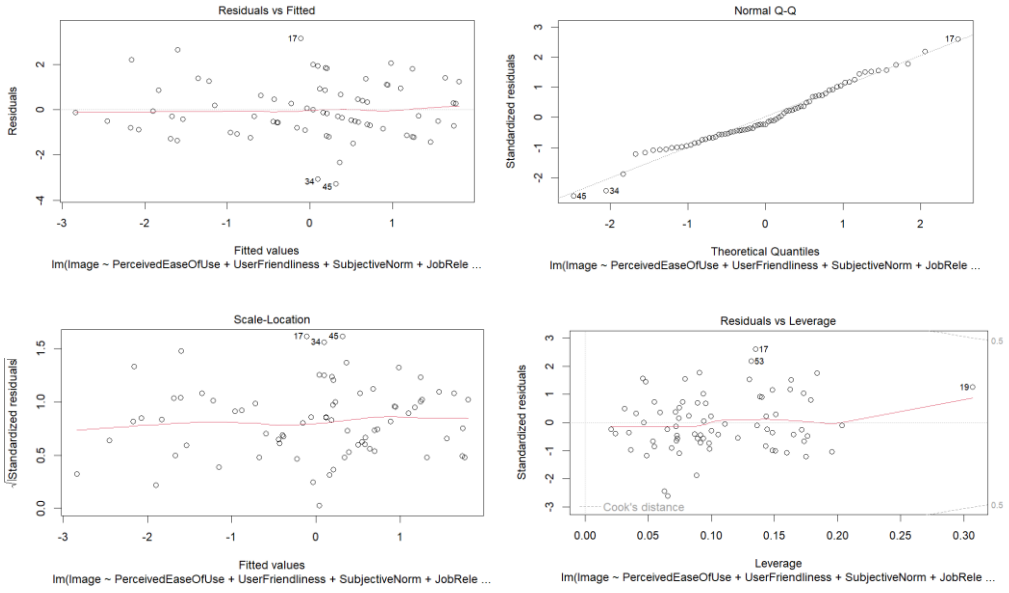
Table F3: Breusch-Pagan test, null-hypothesis of homoscedastic residuals cannot be rejected if p<0.05.

BP	df	p-value
10.504	7	0.1618

Table F4: Shapiro-Wilk normality test, null-hypothesis of normally distributed residuals cannot be rejected if p<0.05.

W	p-value
0.97717	0.1933

Figure F1: Residuals visualizations, OLS Model, Mediation of other variables through Image (IMG), full sample



Appendix G: Sobel test of mediation effects of SN and JR through IMG on PU-F.

	Sobel	Aroian	Goodman
z.value	2.861411888	2.834343992	2.889270390
p.value	0.004217587	0.004591987	0.003861369

Sobel test, mediation effect of JR through IMG on PU-F, full sample. Indicates a significant mediation effect if $p<0.05$.

	Sobel	Aroian	Goodman
z.value	3.215422495	3.179606419	3.252476825
p.value	0.001302527	0.001474752	0.001144039

Sobel test, mediation effect of SN through IMG on PU-F, full sample. Indicates a significant mediation effect if $p<0.05$.

Appendix H: Secondary testing through Bootstrapping

Table H1: Bootstrap (BCa, 10,000 samples) of SN, JR, PEOU, UF, QUAL, RES-F, and FAM on IMG:

	SN on IMG	JR on IMG	PEOU on IMG	UF on IMG	QUAL on IMG	RES-F on IMG	FAM on IMG	IMG~ ~IMG
Min.:	-0.3684	-0.3221	-0.59041	-0.45225	-0.4369	-0.42234	-0.434211	0.5876
1st Quantile:	0.229	0.2889	-0.16808	0.03061	0.1058	-0.04763	-0.0716061	1.1646
Median:	0.3263	0.3836	-0.08115	0.12946	0.2059	0.03905	0.0002483	1.327
Mean:	0.3238	0.3841	-0.07673	0.13009	0.2092	0.0376	-0.000604	1.3393
3rd Quantile:	0.4227	0.4797	0.01147	0.22962	0.3115	0.12249	0.0734383	1.502
Max.	0.9988	0.9285	0.48444	0.73126	0.8037	0.55254	0.4833865	2.3385

Table H2: Confidence interval of bootstrapped indirect effect, checking for zero inclusion:

Variable	CI 2.5%	CI 97.5%	Significant (no zero inclusion)
SN	0.03168299	0.59932156	**Yes**
JR	0.1050211	0.6656667	**Yes**
PEOU	-0.3294936	0.1945649	No
UF	-0.1558455	0.4232284	No
QUAL	0.1050211	0.6656667	**Yes**
RES-F	-0.09083205	0.51863604	No
FAM	-0.2112087	0.2847222	No

Table H3: Bootstrap (BCa, 10,000 samples) of IMG, SN, JR, PEOU, UF, QUAL, RES-F, and FAM on PU-F:

	IMG on PU-F	SN on PU-F	JR on PU-F	PEOU on PU-F	UF on PU-F	QUAL on PU-F	RES-F on PU-F	FAM on PU-F	PU-F~ ~PU-F
Min.:	-0.0172	-0.377958	-0.36318	-0.446499	-0.45471	-0.08105	0.1625	-0.36788	0.2165
1st Quantile:	0.1939	-0.112683	-0.11581	-0.007392	-0.11063	0.21153	0.4965	-0.04713	0.4189
Median:	0.2386	-0.061735	-0.063474	0.056809	-0.04928	0.2663	0.5463	0.01397	0.4691
Mean:	0.2395	-0.060808	-0.061933	0.054292	-0.05007	0.26678	0.545	0.01381	0.468
3rd Quantile:	0.2828	-0.008731	-0.009047	0.118244	0.01189	0.32086	0.5961	0.0741	0.5175
Max.	0.5972	0.281753	0.308042	0.380653	0.30035	0.64211	0.8027	0.38419	0.7267

Table H3: Confidence interval of bootstrapped effect, checking for zero inclusion:

Variable	CI 2.5%	CI 97.5%	Significant (no zero inclusion)
SN through IMG	0.007266529	0.158814945	**Yes**
JR through IMG	0.01773433	0.19570369	**Yes**
IMG	0.1047099	0.3793718	**Yes**
SN	-0.21602440	0.09556196	No
JR	-0.21844930	0.09846183	No
PEOU	-0.1337583	0.2310860	No
UF	-0.2319167	0.1280155	No
QUAL	-0.21844930	0.09846183	No
RES-F	0.1084316	0.4342271	**Yes**
FAM	0.3924013	0.6868698	**Yes**

Appendix I: Modelling Perceived Usefulness for Students (PU-S), full sample (N=75)

Table I1: OLS Model, Perceived Usefulness for Students (PU-S), full sample (N=75)

PU-S	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-6.121e-17	8.916e-02	0.000	1.0000
PEOU-S	5.393e-02	6.098e-02	0.885	0.3796
UF	8.147e-02	9.198e-02	0.886	0.3790
SN	-4.902e-02	7.990e-02	-0.614	0.5416
IMG	3.203e-01	7.356e-02	4.354	4.76e-05 ***
JR	-8.873e-02	8.293e-02	-1.070	0.2886
QUAL	2.090e-01	8.214e-02	2.544	0.0133 *
RES-S	4.712e-01	8.516e-02	5.534	5.81e-07 ***
FAM	1.337e-01	7.930e-02	1.686	0.0965 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7722 on 66 degrees of freedom
Multiple R-squared: 0.7335, Adjusted R-squared: 0.7012
F-statistic: 22.7 on 8 and 66 DF, p-value: 3.258e-16

Table I2: Variance Inflation Factors, Perceived Usefulness for Students (PU-S), full sample

Variable	VIF
PEOU-S	1.395541
UF	1.679241
SN	1.408372
IMG	1.877763
JR	2.292700
QUAL	1.708544
RES-S	1.881125
FAM	1.536331

Residual analysis, Perceived Usefulness for Students (PU-S), full sample

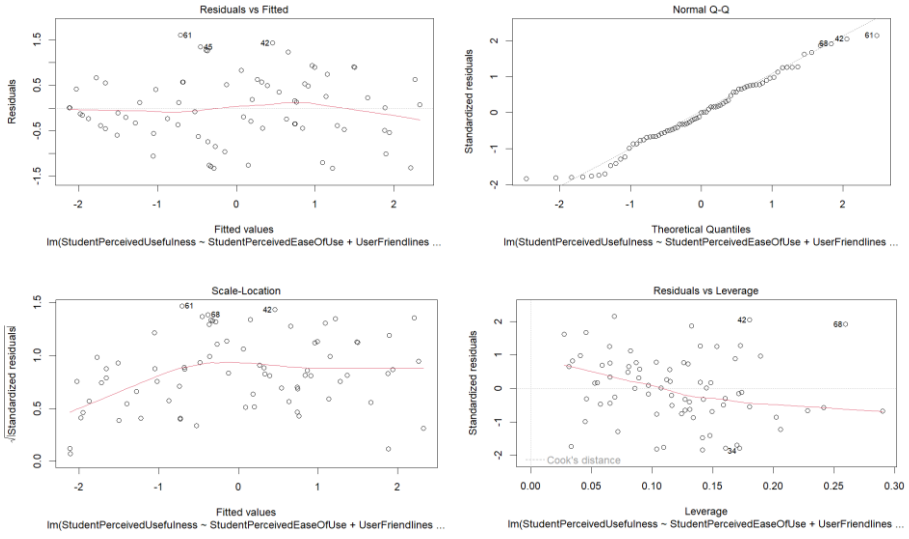
Table I3: Breusch-Pagan test, null-hypothesis of homoscedastic residuals cannot be rejected if p<0.05.

BP	df	p-value
12.467	8	0.1316

Table I4: Shapiro-Wilk normality test, null-hypothesis of normally distributed residuals cannot be rejected if p<0.05.

W	p-value
0.97927	0.2565

Figure I1: Residuals visualizations, OLS Model, Perceived Usefulness for Students (PU-S), full sample



Appendix J: Modelling Perceived Usefulness for Faculty (PU-F), subsample: ChatGPT users only (N=60)

Table J1: OLS Model, Perceived Usefulness (PEOU), ChatGPT users only (N=60)

PU-F	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.05697	0.10747	-0.530	0.59834
PEOU	0.09519	0.08998	1.058	0.29508
UF	-0.08188	0.10708	-0.765	0.44802
SN	-0.08732	0.08499	-1.027	0.30904
IMG	0.24778	0.07692	3.221	0.00222 **
JR	-0.06948	0.09530	-0.729	0.46929
QUAL	0.27936	0.08656	3.227	0.00219 **
RES-F	0.57462	0.07391	7.775	3.27e-10 ***
FAM	0.07857	0.12357	0.636	0.52771

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7487 on 51 degrees of freedom
Multiple R-squared: 0.7851, Adjusted R-squared: 0.7514
F-statistic: 23.3 on 8 and 51 DF, p-value: 1.612e-14

Table J2: Variance Inflation Factors, Perceived Usefulness for Faculty (PU-F), ChatGPT users only

Variable	VIF
PEOU	1.411728
UF	1.562956
SN	1.284070
IMG	1.690712
JR	2.381480
QUAL	1.750904
RES-F	1.541614
FAM	1.414749

Residual analysis, Perceived Usefulness for Faculty (PU-F), ChatGPT users only

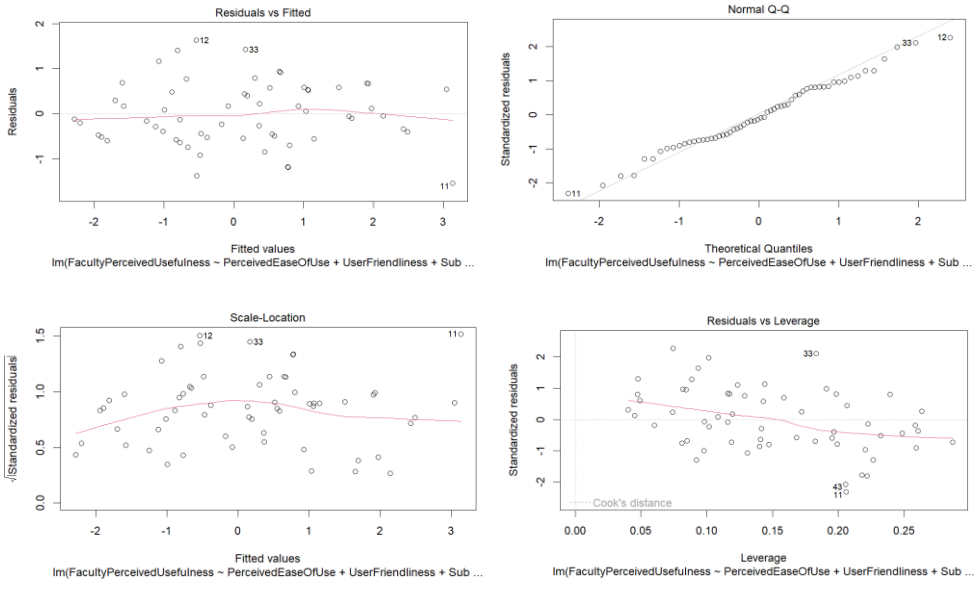
Table J3: Breusch-Pagan test, null-hypothesis of homoscedastic residuals cannot be rejected if p<0.05.

BP	df	p-value
9.0726	8	0.3362

Table J4: Shapiro-Wilk normality test, null-hypothesis of normally distributed residuals cannot be rejected if p<0.05.

W	p-value
0.98765	0.8048

Figure J1: Residuals visualizations, OLS Model, Perceived Usefulness (PU-F), ChatGPT users only



Appendix K: Modelling Mediation of other variables through Image (IMG) on PU-F, subsample: ChatGPT users only (N=60)

Table K1: OLS Model, Mediation of other variables through Image (IMG), ChatGPT users only (N=60)

Image	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.19309	0.19190	1.006	0.3190
PEOU	0.06224	0.16200	0.384	0.7024
UF	0.03328	0.19301	0.172	0.8638
SN	0.34113	0.14574	2.341	0.0231 *
JR	0.36538	0.16417	2.226	0.0304 *
QUAL	0.28000	0.15115	1.852	0.0696 .
RES-F	-0.04281	0.13312	-0.322	0.7491
FAM	-0.08763	0.22245	-0.394	0.6953

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.35 on 52 degrees of freedom
Multiple R-squared: 0.4085, Adjusted R-squared: 0.3289
F-statistic: 5.131 on 7 and 52 DF, p-value: 0.0001751

Table K2: Variance Inflation Factors, Mediation of other variables through Image (IMG), ChatGPT users only

Variable	VIF
PEOU	1.407732
UF	1.562063
SN	1.161672
JR	2.174361
QUAL	1.642515
RES-F	1.538554
FAM	1.410540

Residual analysis, Mediation of other variables through Image (IMG), ChatGPT users only

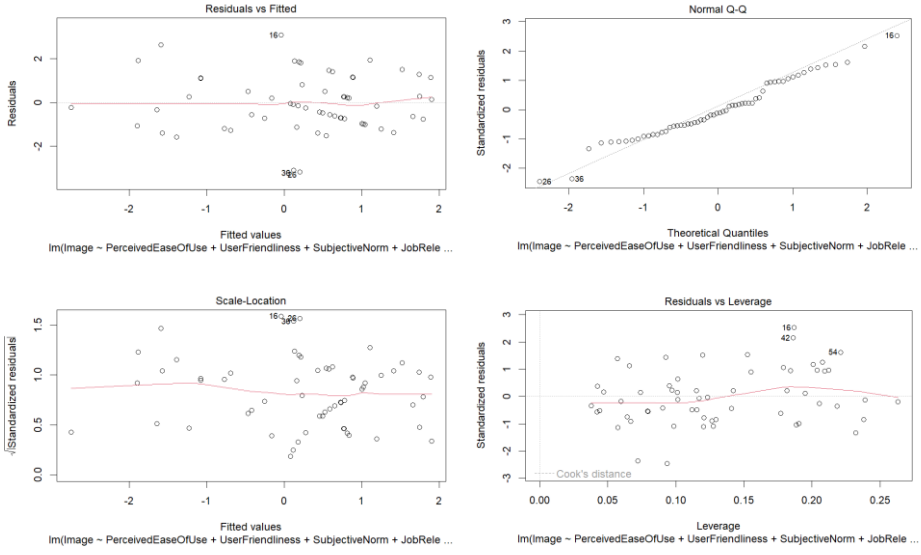
Table K3: Breusch-Pagan test, null-hypothesis of homoscedastic residuals cannot be rejected if p<0.05.

BP	df	p-value
17.067	7	0.01697

Table K4: Shapiro-Wilk normality test, null-hypothesis of normally distributed residuals cannot be rejected if p<0.05.

W	p-value
0.97543	0.2664

Figure K1: Residuals visualizations, OLS Model, Mediation of other variables through Image (IMG), ChatGPT users only



Appendix M: Modelling Perceived Usefulness for Students (PU-S), subsample: ChatGPT users only (N=60)

Table M1: OLS Model, Perceived Usefulness for Students (PU-S), ChatGPT users only (N=60)

PU-S	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.01462	0.12296	0.119	0.9058
PEOU-S	0.05168	0.07242	0.714	0.4788
UF	0.05068	0.12032	0.421	0.6754
SN	-0.02828	0.09707	-0.291	0.7719
IMG	0.31438	0.08914	3.527	0.0009 ***
JR	-0.09325	0.10973	-0.850	0.3994
QUAL	0.23246	0.10115	2.298	0.0257 *
RES-S	0.47944	0.10141	4.728	1.83e-05 ***
FAM	0.14385	0.13998	1.028	0.3089

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8566 on 51 degrees of freedom
Multiple R-squared: 0.7034, Adjusted R-squared: 0.6568
F-statistic: 15.12 on 8 and 51 DF, p-value: 4.396e-11

Table M2: Variance Inflation Factors, Perceived Usefulness for Students (PU-S), ChatGPT users only

Variable	VIF
PEOU-S	1.365365
UF	1.507539
SN	1.279799
IMG	1.734966
JR	2.412264
QUAL	1.826424
RES-S	1.835928
FAM	1.386887

Residual analysis, Perceived Usefulness for Students (PU-S), ChatGPT users only

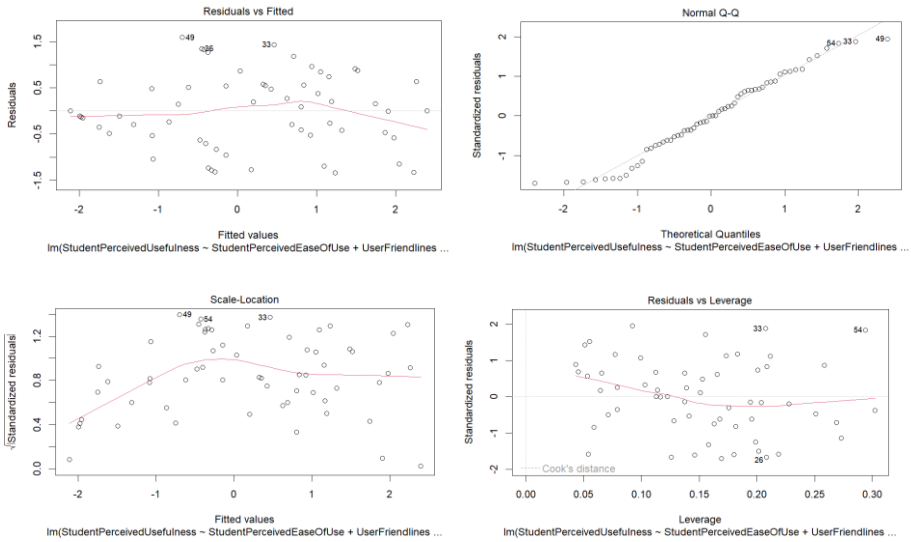
Table M3: Breusch-Pagan test, null-hypothesis of homoscedastic residuals cannot be rejected if p<0.05.

BP	df	p-value
7.8514	8	0.4481

Table M4: Shapiro-Wilk normality test, null-hypothesis of normally distributed residuals cannot be rejected if p<0.05.

W	p-value
0.97066	0.1573

Figure M1: Residuals visualizations, OLS Model, Perceived Usefulness for Students (PU-S), ChatGPT users only



Appendix N: Modelling Perceived Usefulness for Faculty (PU-F), subsample: faculty members only (N=43)

Table N1: OLS Model, Perceived Usefulness (PEOU), faculty members only (N=43)

PU-F	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	7.502e-17	1.249e-01	0.000	1.00000
PEOU	8.944e-02	1.091e-01	0.820	0.418010
UF	-6.965e-02	1.324e-01	-0.526	0.602181
SN	1.231e-02	1.061e-01	0.116	0.908289
IMG	2.895e-01	1.088e-01	2.661	0.011805 *
JR	-7.192e-02	1.166e-01	-0.617	0.541619
QUAL	3.597e-01	1.138e-01	3.162	0.003286 **
RES-F	4.325e-01	1.017e-01	4.253	0.000156 ***
FAM	8.620e-02	1.192e-01	0.723	0.474453

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8191 on 34 degrees of freedom
Multiple R-squared: 0.7629, Adjusted R-squared: 0.7071
F-statistic: 13.67 on 8 and 34 DF, p-value: 1.263e-08

Table N2: Variance Inflation Factors, Perceived Usefulness for Faculty (PU-F), faculty members only

Variable	VIF
PEOU	1.291246
UF	1.580162
SN	1.271013
IMG	2.251740
JR	2.240133
QUAL	1.589638
RES-F	1.620546
FAM	1.300899

Residual analysis, Perceived Usefulness for Faculty (PU-F), faculty members only

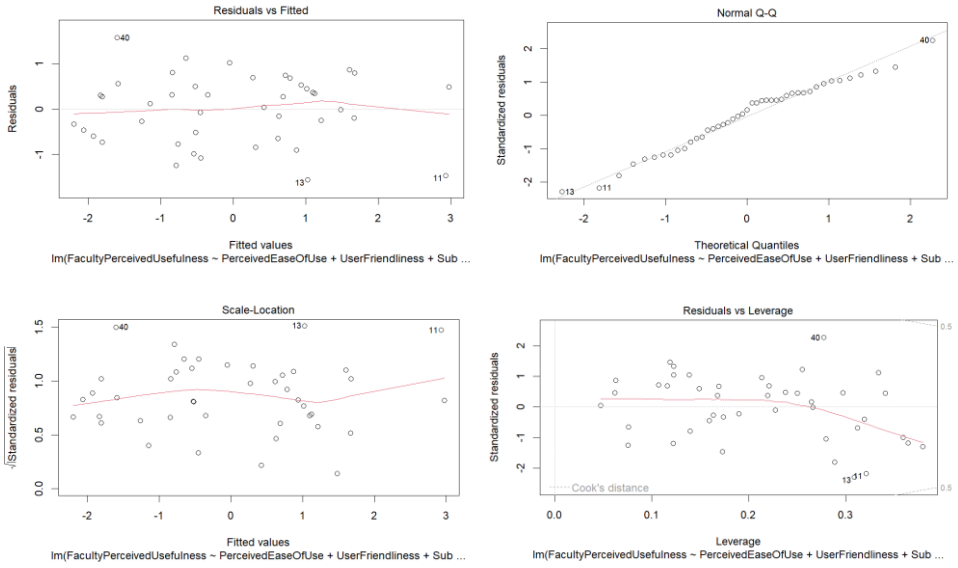
Table N3: Breusch-Pagan test, null-hypothesis of homoscedastic residuals cannot be rejected if p<0.05.

BP	df	p-value
5.0188	8	0.7556

Table N4: Shapiro-Wilk normality test, null-hypothesis of normally distributed residuals cannot be rejected if p<0.05.

W	p-value
0.97872	0.5979

Figure N1: Residuals visualizations, OLS Model, Perceived Usefulness (PU-F), faculty members only



Appendix O: Modelling Mediation of other variables through Image (IMG) on PU-F, subsample: faculty members only (N=43)

Table O1: OLS Model, Mediation of other variables through Image (IMG), faculty members only (N=43)

Image	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-1.683e-16	1.941e-01	0.000	1.00000
PEOU	1.027e-01	1.686e-01	0.609	0.54649
UF	1.190e-01	2.047e-01	0.581	0.56478
SN	1.957e-01	1.615e-01	1.211	0.23382
JR	4.858e-01	1.616e-01	3.007	0.00486 **
QUAL	2.310e-01	1.724e-01	1.340	0.18900
RES-F	1.029e-02	1.570e-01	0.655	0.51677
FAM	3.082e-03	1.851e-01	0.167	0.86872

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.273 on 35 degrees of freedom
Multiple R-squared: 0.5559, Adjusted R-squared: 0.4671
F-statistic: 6.259 on 7 and 35 DF, p-value: 8.501e-05

Table O2: Variance Inflation Factors, Mediation of other variables through Image (IMG), faculty members only

Variable	VIF
PEOU	1.277708
UF	1.565053
SN	1.219859
JR	1.780326
QUAL	1.512105
RES-F	1.600923
FAM	1.299870

Residual analysis, Mediation of other variables through Image (IMG) faculty members only

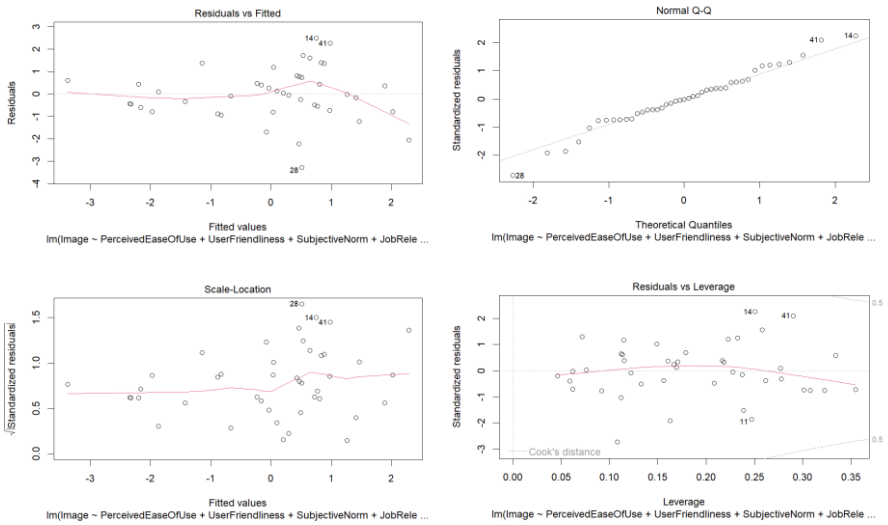
Table O3: Breusch-Pagan test, null-hypothesis of homoscedastic residuals cannot be rejected if p<0.05.

BP	df	p-value
8.567	7	0.2853

Table O4: Shapiro-Wilk normality test, null-hypothesis of normally distributed residuals cannot be rejected if p<0.05.

W	p-value
0.98169	0.7141

Figure O1: Residuals visualizations, OLS Model, Mediation of other variables through Image (IMG), faculty members only



Appendix P: Modelling Perceived Usefulness for Students (PU-S), subsample: faculty members only (N=43)

Table P1: OLS Model, Perceived Usefulness for Students (PU-S), faculty members only (N=43)

PU-S	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-4.291e-17	1.132e-01	0.000	1.000000
PEOU-S	-1.133e-02	8.407e-02	-0.135	0.893573
UF	5.899e-02	1.268e-01	0.465	0.644843
SN	-3.692e-02	9.896e-02	-0.373	0.711375
IMG	3.595e-01	9.914e-02	3.626	0.000931 ***
JR	-5.252e-02	1.082e-01	-0.485	0.630590
QUAL	2.171e-01	1.079e-01	2.012	0.052208 .
RES-S	4.688e-01	1.085e-01	4.322	0.000128 ***
FAM	1.853e-01	1.080e-01	1.715	0.095396 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7426 on 34 degrees of freedom
Multiple R-squared: 0.7951, Adjusted R-squared: 0.7469
F-statistic: 16.49 on 8 and 34 DF, p-value: 1.182e-09

Table P2: Variance Inflation Factors, Perceived Usefulness for Students (PU-S), faculty members only

Variable	VIF
PEOU-S	1.377201
UF	1.765486
SN	1.345583
IMG	2.275757
JR	2.346372
QUAL	1.740175
RES-S	1.732029
FAM	1.300134

Residual analysis, Perceived Usefulness for Students (PU-S), faculty members only

Table P3: Breusch-Pagan test, null-hypothesis of homoscedastic residuals cannot be rejected if p<0.05.

BP	df	p-value
7.5635	8	0.4772

Table P4: Shapiro-Wilk normality test, null-hypothesis of normally distributed residuals cannot be rejected if p<0.05.

W	p-value
0.98286	0.7597

Figure P1: Residuals visualizations, OLS Model, Perceived Ease of Use (PEOU), faculty members only