*education sciences*

MDPI

*Article*

# Perceptions of Project-Based Learning Among Diverse Non-STEM Students: The Role of Quantitative Preparedness and ADHD Status

**Anna Khalemsky [1,*] and Yelena Stukalin [2]**

[1] School of Management, Jerusalem Multidisciplinary College, Jerusalem 91010, Israel
[2] Statistics Unit, The Academic College of Tel Aviv-Yaffo, Tel-Aviv-Yaffo 6818211, Israel; elena@mta.ac.il
* Correspondence: anaha@edu.jmc.ac.il

**Abstract**

Project-based learning (PBL) is increasingly used in higher education to promote authentic, collaborative, and applied learning experiences. However, limited evidence exists regarding how students with varying levels of quantitative preparedness and ADHD experience PBL in introductory analytics courses. This study examined students' perceptions of PBL in introductory statistics and data mining courses taught within non-STEM undergraduate programs. Survey data were collected from 414 students and analyzed using exploratory PCA, reliability analysis, descriptive statistics, and two-way multivariate analysis of variance. Subsequent analyses focused on two theoretically derived composite dimensions: Engagement and Value, and Autonomy. Quantitative preparedness was significantly associated with students' perceptions of PBL, whereas ADHD status alone was not associated with either dimension. Students with lower quantitative preparedness reported higher levels of engagement and perceived value than their more highly prepared peers. An exploratory interaction between quantitative preparedness and ADHD status was observed for the Engagement and Value dimension. The findings suggest that structured project milestones, authentic datasets, collaboration, and instructor guidance may help foster more inclusive analytics learning environments and support meaningful participation among diverse learners. These findings should be interpreted cautiously because they are based on students' self-reported perceptions within a single institution.

**Keywords:** project-based learning; inclusive education; quantitative preparedness; ADHD; statistics education; data mining education



Academic Editors: Geraldine Townend and Marie McGregor

Received: 15 July 2026
Revised: 6 September 2026
Accepted: 8 September 2026
Published: 14 September 2026



## 1. Introduction

The rapid expansion of statistics, data science, and analytics education across higher education has created new instructional challenges for lecturers working with increasingly diverse student populations. Introductory analytics-related courses are now taught not only within STEM disciplines, but also across management, social sciences, health sciences, and other non-technical programs (Marks & AL-Ali, 2022; McLaren et al., 2022). As a result, classrooms often include students with highly heterogeneous quantitative preparedness, varying levels of engagement and perceived value, and different learning profiles (Khalemsky & Stukalin, 2024). While some students enter higher education with strong prior exposure to quantitative reasoning, others begin introductory analytics courses with limited preparedness and substantial anxiety toward quantitative learning (Binder

 https://doi.org/10.3390/educsci16091503

et al., 2019; Paechter et al., 2017; Schneider & Simonsmeier, 2025). These differences create substantial challenges for instructors attempting to maintain engagement and support meaningful learning in introductory courses.

Traditional lecture-based instruction in quantitative courses has often been criticized for emphasizing procedural execution rather than conceptual understanding and authentic reasoning with data (Garfield & Ben-Zvi, 2009; MacGillivray & Pereira-Mendoza, 2011). Such approaches may disadvantage students with weaker quantitative preparedness, particularly in introductory courses where students are simultaneously expected to acquire statistical concepts, interpret findings, and use unfamiliar analytical software tools. In response, higher education increasingly incorporates active learning approaches intended to improve engagement, participation, and applied understanding (Auster & Wylie, 2006; Blumenfeld et al., 1991). These approaches include collaborative learning, inquiry-based learning, experiential learning, formative assessment strategies, and project-based learning (PBL) (Auster & Wylie, 2006; Blumenfeld et al., 1991; Fernandes, 2014). The broader transition toward student-centered learning environments has accelerated further alongside the integration of digital technologies, hybrid learning models, and applied data science education in higher education (Marks & AL-Ali, 2022). PBL was selected because it simultaneously incorporates autonomy, collaboration, authentic datasets, and extended analytical decision-making.

Among these approaches, PBL has attracted growing attention in analytics and quantitative education because it emphasizes authentic problem-solving, collaboration, and applied reasoning with real-world data (Fernandes, 2014; MacGillivray & Pereira-Mendoza, 2011). In statistics and data mining courses, PBL may enable students to work with datasets, formulate research questions, select analytical methods, interpret findings, and communicate conclusions in contexts that resemble real analytical practice (Ben-Zvi et al., 2018; Garfield & Ben-Zvi, 2009; Gould, 2021). Unlike traditional instructional models that focus primarily on procedural fluency, PBL is intended to encourage students to engage actively in the analytical process and to connect quantitative methods with practical decision-making (Blumenfeld et al., 1991; Karaçalli & Korur, 2014). Previous studies have also suggested that project-based environments may enhance strategic thinking, communication skills, and learner autonomy while helping students develop engagement in quantitative tasks (Eliyasni et al., 2019; Saenab et al., 2018).

At the same time, the implementation of PBL in heterogeneous higher education classrooms presents important pedagogical challenges. Prior quantitative preparedness may influence how students engage with analytical tasks, interact with peers, and perceive their ability to succeed in project-based environments (Binder et al., 2019; Schneider & Simonsmeier, 2025). Students entering introductory analytics courses with limited quantitative preparedness may require additional structure, scaffolding, and instructor guidance in order to participate meaningfully in open-ended analytical activities (Paechter et al., 2017). Furthermore, collaborative project work may create uneven participation patterns, where some students dominate the analytical process while others contribute less actively (Wing-yi Cheng et al., 2008). PBL also requires substantial instructor involvement in designing projects, providing feedback, and supporting students throughout extended analytical tasks (Kokotsaki et al., 2016; Pilot et al., 2023).

Introductory analytics courses differ from many other higher education subjects because students must simultaneously develop conceptual understanding, interpret data, operate analytical software, and make methodological decisions (De Veaux et al., 2017). These multiple demands often create substantial cognitive load for students with limited quantitative preparation. Consequently, identifying instructional approaches that support participation across diverse learner populations represents a particularly important challenge in analytics education.

An additional challenge concerns students with Attention-Deficit/Hyperactivity Disorder (ADHD) and related attention-regulation difficulties. ADHD is increasingly recognized among students in higher education and is associated with difficulties in sustained attention, working memory, organization, and self-regulation (Álvarez-Godos et al., 2023; Emmers et al., 2017; Frazier et al., 2007). These challenges may become particularly pronounced in introductory quantitative courses that require sequential reasoning, interpretation of abstract concepts, and sustained engagement with analytical software and multi-step procedures (DuPaul et al., 2009; Hartung et al., 2022; Weyandt & DuPaul, 2006). Students with ADHD may also experience difficulty organizing complex tasks and managing long-term assignments, especially in traditional lecture-centered learning environments (Allsopp et al., 2005; Hartung et al., 2022). At the same time, previous research suggests that learning environments emphasizing autonomy, practical application, collaboration, and structured support may help sustain engagement among students with attention-related challenges (Emmers et al., 2017; Hartung et al., 2022).

PBL may therefore provide a particularly relevant framework for inclusive analytics education. By dividing complex assignments into smaller stages, encouraging collaboration, and allowing students to work with meaningful datasets, PBL may help to reduce some of the barriers associated with low quantitative engagement, perceived value and executive-function difficulties. Structured project milestones, instructor feedback, and opportunities for iterative learning may help students remain engaged even when they initially perceive themselves as weak quantitative learners. However, despite growing interest in PBL and inclusive higher education practices, limited evidence exists regarding how diverse learner characteristics shape students' experiences in project-based introductory analytics courses, particularly in relation to prior quantitative preparedness and ADHD status.

Although project-based learning has been widely studied in statistics, STEM, and higher education contexts, considerably less attention has been devoted to understanding how diverse learner characteristics influence students' experiences within project-based environments. In particular, limited evidence exists regarding the combined role of prior quantitative preparedness and ADHD status in introductory analytics courses. Most previous studies have focused on academic achievement or general satisfaction, rather than examining how different groups of students perceive engagement, autonomy, support, and the value of project-based learning. This gap is particularly important in non-STEM programs, where students often enter quantitative courses with highly heterogeneous educational backgrounds.

Accordingly, this study examines students' perceptions of project-based learning in introductory statistics and data mining courses taught in non-STEM higher education programs. Specifically, the study focuses on the role of prior quantitative preparedness and ADHD status in shaping students' engagement, perceptions of support, autonomy, and perceived value of project-based learning environments. The study addresses the following research questions:

RQ1: How does prior quantitative preparedness relate to students' perceptions of project-based learning in introductory analytics courses?

RQ2: How does ADHD status relate to students' experiences in project-based learning environments?

RQ3: Does prior quantitative preparedness interact with ADHD status in shaping students' perceptions of project-based learning?

The present study contributes to the literature in several ways. First, it extends research on project-based learning to introductory statistics and data mining courses taught within

https://doi.org/10.3390/educsci16091503

non-STEM higher education programs. Second, it examines how quantitative preparedness and ADHD status relate to students' perceptions of engagement, autonomy, and the perceived value of project-based learning. Third, it challenges common assumptions that students with weaker quantitative preparedness necessarily experience lower engagement in quantitative courses. Finally, the study provides practical evidence regarding the design of inclusive project-based learning environments capable of supporting students with diverse educational backgrounds and learning profiles.

## 2. Supporting Diverse Learners in Analytics Education

The expansion of analytics, statistics, and data science education across higher education has intensified the need for instructional approaches capable of supporting increasingly heterogeneous student populations. Introductory analytics courses often include students with substantial differences in prior quantitative preparedness, engagement, learning strategies, and self-regulation skills (Binder et al., 2019; Schneider & Simonsmeier, 2025). These differences are particularly important in non-STEM programs, where students may have limited previous exposure to quantitative reasoning and analytical software tools. As analytics education increasingly emphasizes applied problem-solving, interpretation of findings, and data-driven decision-making, instructors are required to design learning environments that remain accessible and engaging for learners with diverse academic and cognitive profiles (Ben-Zvi et al., 2018; Gould, 2021; McLaren et al., 2022).

Prior quantitative preparedness plays an important role in shaping students' experiences in introductory quantitative courses. Students entering higher education with stronger quantitative preparedness often demonstrate greater engagement when approaching analytical tasks, whereas students with weaker preparedness may experience uncertainty, anxiety, and reduced participation (Binder et al., 2019; Paechter et al., 2017; Schneider & Simonsmeier, 2025). Previous research has shown that prior knowledge influences not only academic performance but also students' ability to integrate new concepts into existing cognitive frameworks (Schneider & Simonsmeier, 2025). In introductory analytics courses, insufficient preparedness may create difficulties in understanding statistical reasoning, selecting analytical procedures, interpreting outputs, and connecting quantitative findings with practical contexts (Binder et al., 2019).

At the same time, prior quantitative preparedness does not necessarily determine students' capacity to engage meaningfully in analytics learning environments. Research in statistics education suggests that instructional approaches emphasizing interpretation, reasoning, and authentic investigation may help reduce the dominance of purely procedural thinking and encourage broader participation among students with weaker quantitative preparedness (Garfield & Ben-Zvi, 2009; MacGillivray & Pereira-Mendoza, 2011).

An additional dimension of learner diversity concerns students with Attention-Deficit/Hyperactivity Disorder (ADHD) and related attention-regulation difficulties. ADHD is increasingly recognized in higher education and is associated with challenges in sustained attention, executive functioning, working memory, organization, and time management (Álvarez-Godos et al., 2023; Emmers et al., 2017; Frazier et al., 2007). These characteristics may influence students' experiences in quantitative learning environments, particularly in courses requiring sequential reasoning, multi-step problem solving, and prolonged engagement with abstract concepts (DuPaul et al., 2009; Hartung et al., 2022; Weyandt & DuPaul, 2006). Students with ADHD may also experience difficulties in organizing long-term assignments and maintaining sustained focus during traditional lecture-centered instruction (DuPaul et al., 2009; Hartung et al., 2022). Quantitative courses that rely heavily on passive listening, procedural demonstrations, and rigid instructional structures may therefore create additional barriers to participation and engagement. At the same time,

https://doi.org/10.3390/educsci16091503

previous research suggests that instructional approaches incorporating autonomy, practical application, flexibility, and structured support may help sustain motivation and participation among students with attention-related challenges (Allsopp et al., 2005; Emmers et al., 2017; Hartung et al., 2022).

Self-Determination Theory (Deci & Ryan, 2000) suggests that autonomy, competence, and relatedness support intrinsic motivation. PBL may address all three needs by allowing students to make decisions, work collaboratively, and experience progressive mastery of analytical tasks. Within this context, project-based learning offers several characteristics that may support inclusive analytics education. First, PBL enables students to work on authentic analytical problems using real or realistic datasets, thereby increasing perceived relevance and engagement (Blumenfeld et al., 1991; Fernandes, 2014). Second, project-based environments often divide complex assignments into smaller sequential stages, which may help reduce cognitive overload and support task organization (Kokotsaki et al., 2016). Third, collaboration with peers and regular instructor feedback may provide additional academic and emotional support throughout the analytical process (Pilot et al., 2023; Saenab et al., 2018). Finally, allowing students to select project topics or datasets may increase autonomy and ownership of learning, factors associated with stronger intrinsic motivation and persistence (Blumenfeld et al., 1991; Pilot et al., 2023).

The present study builds on these perspectives by examining how prior quantitative preparedness and ADHD status relate to students' perceptions of project-based learning in introductory analytics courses. Rather than focusing solely on academic performance, the study explores students' experiences of engagement, autonomy, collaboration, and perceived learning value within project-based learning environments. Understanding how diverse learners experience PBL may help instructors design more inclusive analytics courses capable of supporting participation across varying educational and cognitive backgrounds.

Taken together, prior research suggests that both quantitative preparedness and ADHD status may shape students' experiences in project-based learning environments. However, the direction and magnitude of these relationships remain unclear, particularly in introductory analytics courses taught to non-STEM students. Accordingly, the present study investigates whether these learner characteristics are associated with students' perceptions of engagement, autonomy, support, and perceived value within project-based learning environments.

## 3. Project-Based Learning in Introductory Analytics Courses

The project-based learning (PBL) model examined in this study was implemented in introductory statistics and data mining courses taught within non-STEM undergraduate programs. The courses were designed to combine quantitative reasoning, applied analytics, and authentic problem-solving through semester-long practical projects. Across all courses, students worked with analytical software tools such as SPSS, STATA, Excel, and, in the data mining courses, additional data mining platforms such as Weka (Karaçalli & Korur, 2014; Khalemsky et al., 2024). The projects emphasized real-world applications, iterative learning, and independent analytical decision-making.

The instructional design of the projects was based on several core pedagogical principles associated with project-based learning. PBL emphasizes active participation, authentic problem-solving, collaboration, and the integration of theoretical knowledge with practical application (Auster & Wylie, 2006; Blumenfeld et al., 1991; Fernandes, 2014). Accordingly, students were required to engage actively in the analytical process rather than passively reproduce predefined procedures. The projects were structured as sequential tasks completed throughout the semester, allowing students to receive ongoing feedback and gradually develop analytical skills. Students were also encouraged to work with datasets and research topics connected to their disciplinary interests, thereby increasing the perceived relevance of quantitative learning and promoting learner autonomy (Blumenfeld et al., 1991; Eliyasni et al., 2019). Throughout the semester, instructors provided continuous guidance, including methodological support, feedback sessions, and assistance with analytical interpretation (Pilot et al., 2023).

The introductory statistics courses focused on helping students develop foundational quantitative reasoning skills through authentic research-oriented projects. Students were required to formulate research questions, construct or select datasets, perform descriptive and inferential analyses, and interpret findings using statistical software tools. The projects integrated topics such as descriptive statistics, correlation analysis, regression analysis, hypothesis testing, and analysis of variance. This instructional approach aligns with recommendations in the statistics education literature emphasizing interpretation, reasoning, and authentic investigation rather than procedural memorization alone (Garfield & Ben-Zvi, 2009; MacGillivray & Pereira-Mendoza, 2011). Students were therefore encouraged not only to execute analyses but also to interpret outputs, justify analytical choices, and connect statistical findings with substantive research questions.

The data mining course emphasized introductory knowledge discovery processes and applied machine learning tasks (Khalemsky et al., 2024). Students worked with real-world datasets containing multiple predictors and a categorical target variable. The projects included stages such as defining an analytical problem, reviewing the relevant literature, preprocessing data, comparing classification algorithms, evaluating model performance, and interpreting clustering solutions. Students were also required to compare multiple classification experiments and evaluate models using indicators such as accuracy, precision, recall, Cohen's kappa, and area under the curve (AUC). Such activities reflect broader trends in analytics education that increasingly emphasize authentic data analysis, experimentation, and applied decision-making (Ben-Zvi et al., 2018; Gould, 2021; McLaren et al., 2022). Figure 1 demonstrates the PBL schematic process.

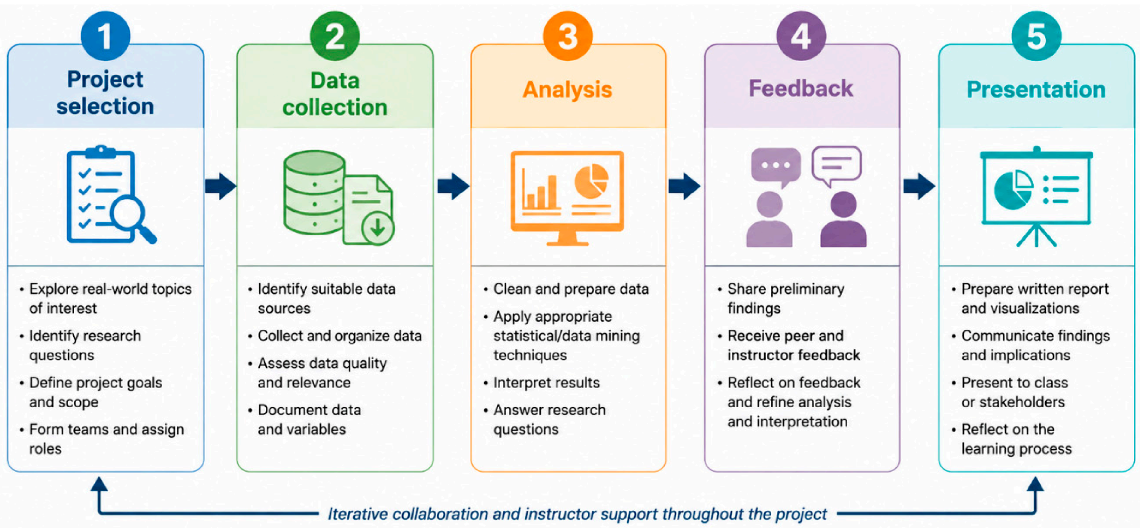


**Figure 1.** PBL process in introductory analytics courses.

Although the statistics and data mining courses differed in analytical complexity and software usage, both learning environments shared several common instructional characteristics. In both contexts, students engaged in extended analytical projects spanning substantial portions of the semester, collaborated with peers, interacted regularly with instructors, and were expected to make independent analytical decisions. These characteristics are consistent with the previous literature suggesting that project-based environments may support strategic thinking, communication skills, learner autonomy, and engagement with complex analytical tasks (Eliyasni et al., 2019; Fernandes, 2014; Saenab et al., 2018).

The instructional approach was particularly relevant in the context of heterogeneous classrooms characteristic of non-STEM higher education programs. Students entered the courses with varying levels of quantitative preparedness and engagement and perceived value, and the project structure was intended to provide incremental scaffolding throughout the learning process. Assignments were divided into smaller stages requiring progressive completion of analytical tasks, allowing students to build engagement gradually while receiving continuous feedback. Previous research suggests that such structured support mechanisms may be particularly beneficial in complex learning environments requiring self-regulation and sustained engagement (Hartung et al., 2022; Kokotsaki et al., 2016; Pilot et al., 2023). The combination of structure, collaboration, and autonomy therefore formed the pedagogical basis for the project-based learning environments examined in the present study.

Detailed examples of project instructions and analytical tasks are presented in Appendix A.

## 4. Methods

### *4.1. Research Design*

This study used a cross-sectional survey design to examine how students' quantitative preparedness and ADHD status relate to their experiences in statistics and data mining courses that utilize Project-Based Learning. The study was not designed to infer causal relationships but to examine associations between background characteristics and students' perceptions of project-based learning. The instrument is included so instructors can adapt items to evaluate PBL implementation in their classrooms.

### *4.2. Participants and Courses*

The study included 414 students from a community college in Israel enrolled in statistics and data mining introductory courses, with 316 participating in statistics courses and 98 in data mining courses. The questionnaire, distributed to approximately 15 student groups during the 2023–2024 academic year (11 learned statistics and 4 learned data mining), comprised three main sections: demographic and personal information, and a practical project questionnaire. The response rate was high due to the questionnaires being administered during frontal lessons—between 60% and 95% of students. Practical projects form a core component of all courses, with students receiving individualized guidance throughout the semester. The projects complete a series of sequential assignments that build progressively on one another.

Statistical courses in non-STEM departments are typically offered in the first year of undergraduate studies as a required component. The statistics curriculum consists of a two-semester course (28–42 academic hours each semester) covering topics such as descriptive statistics, correlation analysis, linear regression, and various statistical inference topics. All courses integrate hands-on work with software tools, and each course includes a practical project that spans the semester. Despite being offered across different departments, the statistics courses share the same core content. The data mining course is offered exclusively within the information systems and data science specializations. It is a mandatory course for these specializations and is taught only in the second academic year. The course opens with an overview of data mining techniques, the process of defining a knowledge problem, and an introduction to analytical tools, classification models, and key clustering approaches (Khalemsky et al., 2024).

### *4.3. Instruments*

#### 4.3.1. Demographic and Personal Information

In addition to questions about gender, age, course type (statistics/data mining), department (social sciences/life sciences/paramedical), academic year, average learning hours per week, employment status, and ADHD report, students were asked questions that emphasized the unique characteristics of the studied population. This included religion (Jewish, Muslim, or other), campus type (main campus with secular Jewish and Muslim students or a Haredi campus where Haredi Jewish female and male students study in separate buildings), and quantitative preparedness level (low, medium, or high). Quantitative preparedness was assessed using students' self-reported level of mathematics completed during high school. In the Israeli secondary education system, mathematics is taught at three nationally standardized levels: 3 units (basic level), 4 units (intermediate level), and 5 units (advanced level). Students were asked: "What was the highest level of mathematics you completed in high school?" Responses were classified as low (3 units), medium (4 units), and high (5 units) quantitative preparedness.

#### 4.3.2. Practical Projects Questionnaire

The questionnaire was developed by the authors based on established theoretical perspectives in project-based learning and higher education, including Self-Determination Theory (Ryan & Deci, 2022), Expectancy-Value Theory (Wigfield & Eccles, 2000) and Social Constructivism (Palincsar, 1998). Together, these frameworks informed items assessing engagement and perceived value, autonomy, collaboration and support, instructional structure, and independent information seeking. The questionnaire consists of 10 Likert-scale questions, rated from 1 to 10, aimed at capturing students' perspectives on various aspects of the practical project. These included the project's importance, interactions with peers and teaching staff, independence in executing the projects, and the potential future application of the knowledge gained. The carefully selected questions were intended to capture nuanced insights into the students' perceptions of PBL in undergraduate non-STEM statistics and data mining courses. The overall content validity index was calculated as the average of the item-level content validity indices across the evaluated criteria. The chosen scale demonstrated high reliability (Cronbach's $\alpha$ = 0.743, content validity index = 0.847). To establish content validity for the newly developed questionnaire, five independent experts were asked to rate each item based on three criteria: relevance, clarity, and coverage of the construct (Lynn, 1986). Using a 4-point scale, we calculated the Item-level Content Validity Index (I-CVI) and the Scale-level Content Validity Index (S-CVI) for each criterion. The results showed moderate relevance (I-CVI = 0.60, S-CVI = 0.84), good clarity (I-CVI = 0.80, S-CVI = 0.92), and moderate coverage (I-CVI = 0.60, S-CVI = 0.88).

### *4.4. Statistical Analysis*

The survey results were analyzed both on the full sample of participants and separately for students in statistics and data mining courses. Statistical analyses were conducted using IBM SPSS Statistics Version 29. Preparedness was self-reported rather than objectively measured.

The complete exploratory PCA analysis results are presented in Supplementary Table S1. Although three components with eigenvalues greater than 1 explained 59.1% of the total variance, several items exhibited substantial cross-loadings (e.g., lecturer support and external information) or relatively low communalities (e.g., teamwork and external information). Consequently, the empirical PCA solution was considered insufficiently interpretable to serve as the sole basis for scale construction. The questionnaire was therefore interpreted primarily on theoretical grounds, resulting in two theoretically derived composite dimensions and separate analyses of the remaining items. Items related to teamwork and support demonstrated insufficient internal consistency ($\alpha = 0.537$) and were therefore analyzed separately rather than combined into a composite scale. The exploratory analysis identified three components with eigenvalues greater than 1, explaining 59.1% of the total variance. However, because the PCA solution lacked clear interpretability and theoretical coherence, it was used to inform rather than determine the final measurement structure.

A two-way multivariate analysis of variance (MANOVA) examined whether students' quantitative preparedness (low, moderate or high level) and ADHD status (self-reported yes/no) were associated with students' perceptions of learning in PBL-based courses. MANOVA was selected because the dependent variables were conceptually related and moderately correlated, and the multivariate model reduces the likelihood of inflated Type I error relative to conducting separate univariate tests. Prior to conducting the MANOVA, the assumptions of multivariate normality, homogeneity of covariance matrices, and linear relationships between the dependent variables were evaluated. Homogeneity of covariance matrices was assessed using Box's M test, while standardized residuals and Q–Q plots were examined to evaluate the distribution of residuals. When multivariate effects were present, follow-up univariate ANOVAs were conducted to identify which dimensions contributed to the overall pattern. Because the multivariate interaction was not statistically significant, the subsequent univariate interaction analyses were considered exploratory and were interpreted cautiously. Pairwise comparisons were performed only to describe significant univariate effects and were adjusted for multiple comparisons using the Bonferroni procedure.

## 5. Results

This section presents descriptive findings regarding student characteristics and perceptions of project-based learning, followed by multivariate analyses examining the roles of quantitative preparedness and ADHD status.

### 5.1. Sample Characteristics

Table 1 summarizes the variable distributions. About three-fourths of the sample are students who took statistics courses; the rest took data mining courses. The last column demonstrates the results of the Chi-square tests for the correlation between the course type and the rest of the categorical variables or ANOVA for all continuous variables.

**Table 1.** Variable distributions.

| Variable | Overall | Statistics Courses | DM Courses | Test Type | *p*-Value |
|---|---|---|---|---|---|
| Number of participants | 414 | 316 | 98 | | |
| Gender (N, %) | | | | Chi-square | 0.002 |
| Female | 300 (72.5%) | 241 (76.3%) | 59 (60.2%) | | |
| Male | 113 (27.3%) | 74 (23.4%) | 39 (39.8%) | | |
| Missing | 1 (0.2%) | | | | |
| Age (mean, std) | 24.112 (5.2) | 23.707 (4.8) | 25.388 (6.3) | ANOVA | <0.001 |
| Religion (N, %) | | | | Chi-square | 0.228 |
| Jewish | 321 (77.5%) | 247 (78.2%) | 74 (75.5%) | | |
| Muslim | 79 (19.1%) | 61 (19.3%) | 18 (18.4%) | | |
| Other | 14 (3.4%) | 8 (2.5%) | 6 (6.1%) | | |
| Campus (N, %) | | | | Chi-square | <0.001 |
| Main | 324 (78.3%) | 246 (77.8%) | 78 (79.6%) | | |
| Haredi women | 72 (17.4%) | 62 (19.6%) | 10 (10.2%) | | |
| Haredi men | 18 (4.3%) | 8 (2.6%) | 10 (10.2%) | | |

https://doi.org/10.3390/educsci16091503

**Table 1.** *Cont.*

| Variable | Overall | Statistics Courses | DM Courses | Test Type | *p*-Value |
|---|---|---|---|---|---|
| Academic year (N, %) | | | | | |
| 1st | 188 (45.4%) | 180 (57%) | 8 (8.2%) | Chi-square | <0.001 |
| 2nd | 209 (50.5%) | 122 (38.6%) | 87 (88.8%) | | |
| 3rd | 17 (4.1%) | 14 (4.4%) | 3 (3.1%) | | |
| Department (N, %) | | | | | |
| Social sciences | 253 (61.1%) | 160 (50.6%) | 93 (94.9%) | Chi-square | <0.001 |
| Life sciences | 64 (15.5%) | 59 (18.7%) | 5 (5.1%) | | |
| Paramedical | 97 (23.4%) | 97 (30.7%) | --- | | |
| Quantitative prep. (N, %) | | | | | |
| Low | 175 (42.3%) | 109 (34.5%) | 66 (67.3%) | Chi-square | <0.001 |
| Medium | 171 (41.3%) | 147 (46.5%) | 24 (24.5%) | | |
| High | 68 (16.4%) | 60 (19%) | 8 (8.2%) | | |
| Working status (N, %) | | | | | |
| Yes | 295 (71.3%) | 214 (67.7%) | 81 (82.7%) | Chi-square | 0.004 |
| No | 119 (28.7%) | 102 (32.3%) | 17 (17.3%) | | |
| ADHD diagnosis self-report (N, %) | | | | | |
| Yes | 90 (21.7%) | 72 (22.8%) | 18 (18.4%) | Chi-square | 0.354 |
| No | 324 (78.3%) | 244 (77.2%) | 80 (81.6%) | | |
| Learning hours (mean, std) | 3.657 (5.8) | 3.697 (6.385) | 3.522 (3.489) | ANOVA | 0.801 |

Statistics and data mining students differed on several background characteristics, including gender, age, academic year, department, quantitative preparedness, and employment status. No significant differences were observed in ADHD prevalence or learning hours.

### 5.2. Practical Projects Survey

The second part of the results section highlights a questionnaire designed to assess students' attitudes toward different aspects of project-based learning. Table 2 displays the findings for each of the ten questions.

**Table 2.** Practical projects survey.

| Question | Mean | SD | CI [95%] |
|---|---|---|---|
| Q1: How much integration of the practical project enriches the course? | 6.33 | 2.8 | [6.1, 6.6] |
| Q2: How much more involved do you feel when there is a practical project in the statistics/data mining course? | 5.85 | 2.9 | [5.6, 6.1] |
| Q3: How much do you prefer a structured course? | 6.38 | 2.7 | [6.1, 6.7] |
| Q4: How much do you value autonomy in selecting the topic and datasets? | 6.73 | 2.9 | [6.5, 7.0] |
| Q5: How much do you value autonomy in decision-making? | 6.61 | 2.8 | [6.3, 6.9] |
| Q6: How much do you enjoy teamwork in a practical project? | 6.89 | 3.01 | [6.6, 7.2] |
| Q7: Are you comfortable reaching out to the lecturer for assistance? | 7.61 | 2.7 | [7.3, 7.9] |
| Q8: Are you comfortable reaching out to other students for assistance and collaboration? | 7.35 | 2.7 | [7.1, 7.6] |
| Q9: How often do you rely on external information sources? | 5.69 | 3.1 | [5.4, 6.0] |
| Q10: Do you believe you will use the tools acquired during the project in your future career? | 5.53 | 2.8 | [5.3, 5.8] |

Students generally reported positive perceptions of project-based learning. The highest-rated items were comfort in seeking assistance from the lecturer (M = 7.61, SD = 2.70) and collaboration with peers (M = 7.35, SD = 2.70). Ratings were also relatively high for teamwork (M = 6.89, SD = 3.01) and autonomy in selecting project topics and making project-related decisions (M = 6.73 and 6.61, respectively). Lower ratings were observed for perceived involvement in the course (M = 5.85, SD = 2.90), use of external information sources (M = 5.69, SD = 3.10), and anticipated future use of the acquired tools (M = 5.53, SD = 2.80).

The exploratory PCA analysis served primarily to evaluate the questionnaire structure. Because the resulting factor solution lacked sufficient interpretability, subsequent analyses were conducted using two theoretically defined composite dimensions with satisfactory internal consistency together with the remaining individual items (KMO = 0.740; Bartlett's $\chi^2(45) = 1233.10$, $p < 0.001$). Because several items demonstrated weak or ambiguous component loadings, subsequent analyses were based on theoretically derived dimensions. Engagement and Value included items assessing course enrichment, involvement, and future applicability of acquired skills ($\alpha = 0.808$; M = 5.90, SD = 2.42). Autonomy included items assessing freedom in topic selection and project-related decision-making ($\alpha = 0.860$; M = 6.67, SD = 2.68). Items related to teamwork and support did not demonstrate sufficient internal consistency ($\alpha = 0.537$) and were therefore analyzed separately. Table 3 presents composite dimensions. Table 4 presents descriptive statistics by group.

**Table 3.** Composite dimensions.

| Dimension | Items | Cronbach's $\alpha$ | Mean | SD |
|---|---|---|---|---|
| Engagement & Value | Q1, Q2, Q10 | 0.808 | 5.90 | 2.42 |
| Autonomy | Q4, Q5 | 0.860 | 6.67 | 2.68 |

**Table 4.** Descriptive statistics by group.

| Quantitative Preparedness | ADHD | N | Engagement & Value Mean (SD) | Autonomy Mean (SD) |
|---|---|---|---|---|
| Low | Yes | 48 | 6.65 (2.1) | 7.49 (2.22) |
| Low | No | 127 | 6.30 (2.43) | 6.93 (2.66) |
| Medium | Yes | 32 | 6.16 (1.95) | 6.67 (2.36) |
| Medium | No | 139 | 5.32 (2.55) | 6.35 (2.92) |
| High | Yes | 10 | 4.30 (1.62) | 5.25 (2.44) |
| High | No | 58 | 5.97 (2.33) | 6.41 (2.53) |

### *5.3. Effects of Quantitative Preparedness and ADHD*

A two-way MANOVA was used to examine whether quantitative preparedness (low, medium, high) and ADHD status (yes/no) were associated with students' experiences in the PBL-based statistics and data mining courses. Preliminary diagnostic analyses supported the use of MANOVA. Box's M test was not statistically significant (Box's M = 5.54, $F(3, 142{,}439.35) = 1.82$, $p = 0.142$), indicating that the assumption of equality of covariance matrices across groups was satisfied. Table 5 summarizes MANOVA results.

**Table 5.** Multivariate tests.

| Effect | Pillai's Trace | F | df | $p$ | Partial $\eta^2$ |
|---|---|---|---|---|---|
| Quantitative preparedness | 0.034 | 3.49 | 4, 816 | 0.008 | 0.017 |
| ADHD | 0.001 | 0.11 | 2, 407 | 0.897 | 0.001 |
| Preparedness × ADHD | 0.020 | 2.02 | 4, 816 | 0.090 | 0.010 |

A two-way MANOVA was conducted to examine the effects of quantitative preparedness and ADHD status on the Engagement and Value and Autonomy dimensions. A significant multivariate effect was observed for quantitative preparedness (Pillai's Trace = 0.034, $F(4, 816) = 3.49$, $p = 0.008$, partial $\eta^2 = 0.017$). Neither ADHD status nor the interaction effect reached statistical significance at the multivariate level. To examine whether differences between the Statistics and Data Mining courses influenced the findings, an additional MANOVA was conducted including course type as an additional fixed factor. Course type demonstrated a significant multivariate effect (Pillai's Trace = 0.030, $F(2, 402) = 6.18$,

$p = 0.002$, partial $\eta^2 = 0.030$), indicating overall differences between the two courses. However, none of the interactions involving course type were statistically significant, including the quantitative preparedness × course interaction ($p = 0.689$), the ADHD × course interaction ($p = 0.247$), and the three-way interaction ($p = 0.670$). Table 6 presents the univariate follow-up analysis.

**Table 6.** Univariate tests.

| Dependent Variable | Effect | F | $p$ | Partial $\eta^2$ |
|---|---|---|---|---|
| Engagement & Value | Quantitative preparedness | 5.67 | 0.004 | 0.027 |
| Engagement & Value | ADHD | 0.21 | 0.644 | 0.001 |
| Engagement & Value | Preparedness × ADHD | 3.60 | 0.028 | 0.017 |
| Autonomy | Quantitative preparedness | 4.52 | 0.011 | 0.022 |
| Autonomy | ADHD | 0.06 | 0.801 | 0.000 |
| Autonomy | Preparedness × ADHD | 1.44 | 0.238 | 0.007 |

Although the overall multivariate interaction between quantitative preparedness and ADHD status was not statistically significant, follow-up univariate analyses revealed an exploratory interaction for the Engagement and Value dimension ($F(2, 408) = 3.60$, $p = 0.028$, partial $\eta^2 = 0.017$). Because this interaction was not supported by the overall multivariate test, it should be interpreted cautiously. Bonferroni-adjusted pairwise comparisons indicated that, among students with high quantitative preparedness, those with ADHD reported significantly lower Engagement and Value than those without ADHD (mean difference = −1.665, 95% CI [−3.264, −0.067], $p = 0.041$). No other pairwise comparisons reached statistical significance. As illustrated in Figure 2, Engagement and Value scores among students with ADHD decreased as quantitative preparedness increased, whereas students without ADHD showed the lowest scores at the medium preparedness level and relatively higher scores at the high preparedness level. Accordingly, the largest difference between ADHD groups was observed among students with high quantitative preparedness. Figure 2 illustrates the interaction between quantitative preparedness and ADHD status for the Engagement and Value dimension. Estimated marginal means with 95% confidence intervals are presented in Supplementary Table S1. Students with ADHD and low quantitative preparedness had the highest estimated Engagement and Value scores (M = 6.65, 95% CI [5.98, 7.33]), whereas students with ADHD and high quantitative preparedness had the lowest scores (M = 4.30, 95% CI [2.82, 5.78]). Bonferroni-adjusted pairwise comparisons indicated that among students with high quantitative preparedness, students with ADHD reported significantly lower Engagement and Value than students without ADHD (mean difference = −1.666, 95% CI [−3.264, −0.067], $p = 0.041$). No other pairwise comparisons reached statistical significance.

Although statistically significant effects were observed for quantitative preparedness, the corresponding partial $\eta^2$ values were small, indicating that quantitative preparedness explained only a limited proportion of the variance in students' perceptions. Consequently, the findings should be interpreted as modest associations rather than strong effects. Thus, while quantitative preparedness is associated with differences in students' perceptions of project-based learning, it explains only a modest proportion of the variance. From a teaching perspective, this suggests that prior preparation is associated with perceptions of project-based learning.

In summary, quantitative preparedness emerged as the primary factor associated with students' perceptions of project-based learning. Students with lower preparedness reported higher levels of engagement and perceived value, while perceptions of autonomy also differed across preparedness levels. ADHD status alone was not associated with either dimension. However, an exploratory interaction between quantitative preparedness and ADHD status was observed for the Engagement and Value dimension. Among students with high quantitative preparedness, those with ADHD reported lower Engagement and Value than those without ADHD, although this finding should be interpreted cautiously because the overall multivariate interaction was not statistically significant.

The estimated marginal means and their 95% confidence intervals are presented in Figure 2 (or Supplementary Table S2), providing an estimate of the precision of the group means.

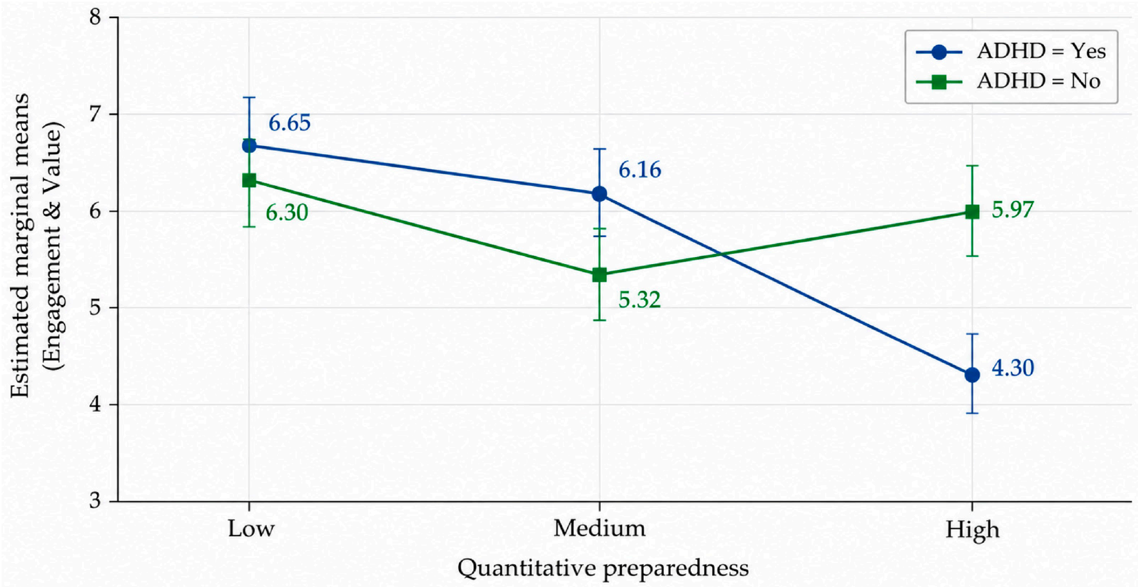


**Figure 2.** Interaction between quantitative preparedness and ADHD status on Engagement and Value.

## 6. Discussion

Project-Based Learning (PBL) is increasingly recognized as an instructional approach that may support active learning, foster autonomy, encourage collaboration, and enable students to integrate course concepts into real analytical processes (Allsopp et al., 2005; Blumenfeld et al., 1991). Unlike lecture-based models that emphasize procedural fluency and compliance with rigid schedules, PBL is inherently flexible and allows instructors to adapt tasks, pace, and expectations to students' personal characteristics and needs (MacGillivray & Pereira-Mendoza, 2011). Previous studies have demonstrated that PBL enhances learner engagement and promotes independence and strategic thinking (e.g., selecting analytical methods, interpreting results, planning next steps) (Garfield & Ben-Zvi, 2009). This makes PBL particularly suitable for culturally diverse classrooms and students with ADHD, who often struggle in traditional formats that require sustained attention, routine, and strict adherence to deadlines (Frazier et al., 2007). When implemented effectively, PBL may encourage meaningful student–lecturer interaction and team collaboration, but achieving these benefits requires instructors to be prepared to provide continuous, personalized guidance throughout the project.

A central finding of this study is that prior quantitative preparedness was significantly associated with students' perceptions of project-based learning. Students with weaker preparedness may perceive larger learning gains because project-based activities provide opportunities to apply concepts in meaningful contexts. In contrast, highly prepared students may enter the course with greater familiarity and therefore experience smaller perceived added value. Students with lower quantitative preparedness reported higher levels of engagement and perceived value, while perceptions of autonomy also differed across preparedness levels. Importantly, students with weaker quantitative preparedness did not report lower engagement or lower perceived benefit (Ryan & Deci, 2022). On the contrary, they often evaluated the project experience more positively than their highly prepared peers.

The observed effect sizes were small, indicating that quantitative preparedness represents only one of many factors associated with students' experiences of project-based learning. Other learner characteristics, instructional practices, and contextual factors are also likely to contribute to engagement and perceived value.

Students with ADHD did not differ significantly from their peers in either Engagement and Value or Autonomy. This finding suggests that project-based learning environments may provide sufficient flexibility and structure to support participation among students with attention-related challenges. The absence of significant negative effects associated with ADHD status is consistent with research suggesting that autonomy, authentic tasks, and ongoing feedback may facilitate engagement among students who experience difficulties in traditional lecture-based environments (Álvarez-Godos et al., 2023; DuPaul et al., 2009).

The exploratory interaction between quantitative preparedness and ADHD status suggests that the association between prior quantitative preparedness and students' perceptions of project-based learning may differ according to ADHD status. Students with ADHD reported lower Engagement and Value as quantitative preparedness increased, whereas students without ADHD showed the lowest Engagement and Value scores at the medium preparedness level and relatively higher scores among highly prepared students. Because the overall multivariate interaction was not statistically significant, this finding should be interpreted cautiously and viewed as hypothesis-generating rather than confirmatory.

The interaction observed for the Engagement and Value dimension should be interpreted cautiously because the overall multivariate interaction was not statistically significant. Nevertheless, the pattern suggests a potentially interesting relationship that may warrant further investigation. In contrast, highly prepared students with ADHD may particularly benefit from the autonomy, flexibility, and applied nature of project-based learning environments, which align with instructional characteristics known to support motivation and sustained engagement among students with attention-related challenges.

Notably, the interaction effect of quantitative preparedness and ADHD status was minimal, yet pedagogically meaningful. Students with ADHD and weaker preparedness did not show compounded disadvantage. Instead, these students remained engaged and persisted throughout the course. This suggests that PBL may disrupt long-held assumptions that success in statistics requires strong quantitative preparedness or neurotypical learning patterns. One possible explanation is that authentic projects may reduce fear of failure, although this mechanism was not directly examined. As suggested by self-determination theory, autonomy, relevance, and competence are key drivers of intrinsic motivation. PBL promotes all three by enabling students to make decisions, explore real datasets, and experience incremental mastery. These findings suggest that project-based learning may help students view quantitative learning as more accessible and relevant to their future studies and professional interests. The observed emphasis on perceived relevance and authentic application is consistent with recent recommendations of the American Statistical Association regarding statistics education.

These findings should be interpreted in light of the study design. Because the analyses were based on cross-sectional self-reported perceptions, they describe associations rather than causal relationships. The findings are particularly relevant for introductory statistics and analytics education, where instructors often assume that students with weaker quantitative preparedness will struggle to engage with analytical tasks. The present findings challenge this assumption. Students with lower preparedness reported equal or higher levels of engagement and perceived value, suggesting that project-based learning may help reduce some of the barriers traditionally associated with quantitative learning. By emphasizing authentic datasets, interpretation, and decision-making rather than procedural calculation alone, PBL may broaden participation in analytics education among students who do not initially perceive themselves as quantitatively strong.

Implications for instructional practice

The findings provide several practical insights for instructors implementing project-based learning in introductory statistics and analytics courses. One possible implication is that structured milestones may be particularly helpful for students with weaker quantitative preparedness. Contrary to common assumptions, these students reported equal or higher levels of engagement and perceived value than their more highly prepared peers. Dividing projects into smaller sequential tasks, providing regular checkpoints, and offering timely feedback may help reduce cognitive overload and enable students to engage with analytical reasoning in a manageable way. However, cognitive load was not directly measured in this study.

Second, autonomy appears to be an important component of successful project-based learning environments. Students reported relatively high levels of appreciation for opportunities to select project topics, choose datasets, and make analytical decisions. Allowing students to work on problems that are personally meaningful may strengthen ownership of the learning process and increase the perceived relevance of quantitative methods.

Third, collaboration and instructor support emerged as highly valued aspects of the learning experience. Students reported particularly positive perceptions of their ability to seek assistance from lecturers and peers throughout the project process. These findings suggest that project-based learning is most effective when autonomy is balanced with accessible support structures, including instructor guidance, peer collaboration, and iterative feedback.

The results also indicate that ADHD status alone was not associated with lower engagement, autonomy, or perceived value. The absence of significant differences between students with and without self-reported ADHD suggests that these groups reported broadly similar perceptions of the PBL experience. Instructors should therefore consider these perceptions alongside additional indicators of learning and achievement when designing inclusive instructional environments. However, this finding should not be interpreted as evidence that PBL effectively addresses ADHD-related learning challenges.

Finally, instructors should be aware that students with strong quantitative preparedness may not automatically perceive greater value in project-based activities. The findings suggest that offering optional advanced analytical tasks for highly prepared students may be worth considering. Providing optional advanced analyses, more complex datasets, additional methodological challenges, or opportunities for independent exploration may help maintain engagement among these students while preserving the inclusive benefits of project-based learning for the broader class.

More broadly, the findings align with contemporary recommendations in statistics and analytics education that emphasize interpretation, communication, problem-solving, and authentic inquiry rather than procedural calculation alone. By engaging students in meaningful analytical tasks using real-world data, project-based learning may help broaden participation in analytics education and support learners with diverse educational backgrounds and learning profiles.

Limitations

Several limitations should be considered when interpreting the findings of this study. First, the study employed a cross-sectional design, preventing causal conclusions regarding the relationships between quantitative preparedness, ADHD status, and students' perceptions of project-based learning. Future longitudinal studies could examine how these perceptions evolve throughout a course and whether project-based learning produces lasting effects on engagement and learning outcomes.

Second, ADHD status was based on self-report rather than clinical documentation. Although self-report measures are commonly used in educational research, some degree of misclassification may have occurred. Similarly, quantitative preparedness was measured through students' self-reported prior quantitative preparedness rather than objective indicators such as mathematics grades, standardized test scores, or prior coursework. Future studies should incorporate objective measures of preparedness to validate and extend the present findings.

Third, the sample was drawn from a single higher education institution and consisted primarily of students enrolled in non-STEM programs. Consequently, the generalizability of the findings to other institutional contexts, countries, disciplines, and student populations may be limited. Replication across multiple institutions and educational settings would strengthen the external validity of the results. In addition, students enrolled in the statistics and data mining courses differed on several background characteristics. Although both courses employed the same project-based learning principles and were analyzed together because they shared a common pedagogical approach, course-specific differences may have influenced students' perceptions and should be examined explicitly in future research.

Fourth, the study relied primarily on self-reported perceptions of engagement, autonomy, support, and perceived value. While these perceptions are important educational outcomes, they do not necessarily reflect actual learning gains, academic performance, or long-term skill development. Future research should combine perception-based measures with objective indicators such as grades, project quality, course completion, retention, and subsequent academic achievement.

Fifth, because most variables were collected through the same survey instrument at a single point in time, the findings may be affected by common-method bias. Future studies could reduce this risk by incorporating multiple data sources, including instructor evaluations, learning analytics, peer assessments, and direct measures of project performance.

Sixth, although the Autonomy dimension demonstrated high internal consistency, it was measured using only two closely related items. Future research should broaden this construct by including additional indicators to strengthen construct validity.

Finally, instructor-related factors were not explicitly examined. Differences in teaching style, feedback practices, project design, and classroom climate may influence students' experiences of project-based learning and could partially explain variation in the observed outcomes. Future studies should investigate how specific instructional practices interact with student characteristics to shape engagement and perceived value in project-based analytics courses.

Despite these limitations, the study provides new evidence regarding the experiences of diverse learners in project-based statistics and data mining courses and offers practical insights for designing more inclusive analytics learning environments. Future studies should examine actual learning outcomes in addition to perceptions.

## 7. Conclusions

This study examined students' perceptions of project-based learning in introductory statistics and data mining courses, with particular attention to the roles of quantitative preparedness and ADHD status. The findings indicate modest but statistically significant associations between quantitative preparedness and students' perceptions of project-based learning. Contrary to common assumptions, students with lower quantitative preparedness often reported higher engagement and perceived value than their more highly prepared peers, suggesting that project-based learning may help reduce barriers frequently associated with introductory quantitative courses.

ADHD status alone was not associated with significant differences in students' perceptions of project-based learning. An exploratory interaction between quantitative preparedness and ADHD status was observed for the Engagement and Value dimension. Although this pattern should be interpreted cautiously because it was not supported by the overall multivariate interaction, it suggests that learner characteristics may influence perceptions of project-based learning in more complex ways than previously assumed. Replication in independent samples is needed. Although the observed effects were modest, these findings highlight the importance of considering individual learner characteristics when designing project-based learning environments.

An exploratory interaction between quantitative preparedness and ADHD status was observed for Engagement and Value, suggesting that learner characteristics may influence perceptions of project-based learning in complex ways. Because this finding was not supported by the overall multivariate interaction, it requires confirmation in future studies.

The study contributes to the growing literature on statistics and analytics education by demonstrating that project-based learning may support meaningful engagement across students with diverse educational backgrounds. The combination of authentic datasets, collaborative work, instructor guidance, and structured project milestones appears particularly valuable for fostering participation in introductory quantitative courses. The study extends existing research by demonstrating that learner characteristics influence perceptions of project-based learning in more complex ways than simple preparedness deficits. In particular, lower preparedness was not associated with lower engagement, challenging assumptions commonly found in quantitative education.

Future research should examine whether these perceptions translate into improved learning outcomes, academic achievement, retention, and long-term development of analytical competencies. Additional studies across institutions, disciplines, and instructional contexts would further clarify how project-based learning can be designed to support increasingly diverse populations of learners in higher education.

**Supplementary Materials:** The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/educsci16091503/s1. Table S1: Principal component analysis of the Practical Projects Questionnaire (N = 414); Table S2: Estimated marginal means (95% confidence intervals) for Engagement and Value by quantitative preparedness and ADHD status.

**Author Contributions:** Conceptualization, A.K. and Y.S.; methodology, A.K. and Y.S.; software, A.K. and Y.S.; validation, A.K. and Y.S.; formal analysis, A.K. and Y.S.; investigation, A.K. and Y.S.; resources, A.K. and Y.S.; data curation, A.K. and Y.S.; writing—original draft preparation, A.K. and Y.S.; writing—review and editing, A.K. and Y.S.; visualization, A.K. and Y.S.; supervision, A.K. and Y.S.; project administration, A.K. and Y.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Ethics Committee of Jerusalem Multidisciplinary College (protocol code 468, 27 November 2023).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The students' feedback de-identified data can be provided upon request.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Appendix A. Practical Projects Instructions for Statistics and Data Mining Courses

**Statistics 1**. Topics learned during the course: descriptive statistics, probability, correlation analysis and linear regression. We adjust the instructions to the type of software we use.

This project simulates the process of conducting data analysis in a real research setting using SPSS version 31, starting from the preparation of the dataset through data processing and drawing conclusions. You are required to work on the project consistently during the first half of the semester and submit a final presentation in the last week.

1. Describe a study that investigates a topic of interest related to your specialization.
2. You may create fictional data, download a dataset from an online repository, or use any other dataset of your choice.
3. Build the dataset in SPSS according to the following guidelines. The dataset should include at least 50 participants and 5 independent variables, all relevant to your research topic (e.g., demographic data, test results before/after treatment, questionnaire responses, blood test indices, etc.). Four independent variables must be quantitative (e.g., age or test scores), and one must be categorical (e.g., gender or medical diagnosis report). The dependent variable (Y) is your choice (e.g., number of treatment sessions, patient satisfaction level, degree of improvement in a specific feature). The dependent variable must be quantitative.
4. Present at least 4 academic articles related to your research topic.
5. Use the APA citation style.
6. Describe the sampling method used in your study. Explain how participants were selected or how you chose your dataset.
7. Formulate research questions and hypotheses.
8. Use graphs to describe the data. You may use pie charts, bar charts, histograms, cumulative relative frequency graphs, scatter plots, etc. Graphs must be created using SPSS. Do not use the same type of graph more than once. You must include at least three different types of graphs. You may also include additional graph types beyond those taught in the course. Discuss what can be inferred from the graphs.
9. Use SPSS descriptive statistics commands to explore the data. What is the distribution of each variable? Are there any outliers? Define an outlier as a value that is at least 2 standard deviations above or below the mean. Based on the "story" you created for your dataset, try to explain the origin of these values.
10. Summarize the research conclusions. What are the key insights from your analysis? How do they relate to your hypotheses and the literature you reviewed?

**Statistics 2**. Topics learned during the course: normal distribution, hypothesis testing, confidence intervals, multiple linear regression.

1. Add an additional categorical variable to your dataset. Run the Crosstabs command in SPSS to calculate Cramér's V between the original categorical variable and the new one you added. Report the value of Cramér's V. Is the correlation statistically significant at the 5% significance level?
2. Run a multiple linear regression in SPSS using all quantitative independent variables. What is the correlation coefficient (R)? Interpret the result. What is the regression model for predicting Y as a function of the independent variables? Which variable is the most influential predictor in your regression model? If you want to keep only the three most influential quantitative predictors, which variables would remain? Explain the meaning of regression coefficients. Create a scatterplot matrix for all quantitative variables in your dataset.

3. Run a Bivariate Correlation analysis in SPSS using only the quantitative variables. Identify which pairs of variables show statistically significant correlations.
4. Choose one quantitative variable. Run a One-Sample T-Test in SPSS at the 5% significance level. Choose a hypothetical population mean ($\mu_0$) appropriate to your data. Test the two-sided hypothesis: $H_1$: $\mu \neq \mu_0$.
5. Add a new column to your dataset that, along with an existing column, creates paired samples. Run a Paired-Samples T-Test in SPSS at the 5% significance level. Can you conclude there is a statistically significant difference between the means?
6. Select one variable from your dataset. Create 15–20 additional observations that appear to belong to a second sample. Run an Independent-Samples T-Test in SPSS at the 5% significance level. Can you conclude that there are statistically significant differences between the groups' means? Repeat the test using a confidence interval for the difference in means.
7. Choose a different variable from your dataset. Create 15–20 additional observations that appear to belong to two other groups (i.e., you now have three groups). Run a One-Way ANOVA in SPSS at the 5% significance level, including post hoc analysis. Are there statistically significant differences between the three means at $\alpha$ = 0.05? Which means do significant differences exist between? Check at the 5% level.

**Data Mining**. Topics learned during the course: introduction to knowledge discovery and data mining, classification analysis, cluster analysis.

1. Task 0 (10%). Search the Kaggle data repository for three datasets related to a topic of interest. Each dataset must include at least 3 independent variables and one categorical dependent variable.
2. Task 1 (40%). Problem Definition and Knowledge Problem Mapping.
3. Slide 1: Problem Definition. Describe the approved topic. Explain why the problem is important. Include a link to the dataset source. Slide 2: Preliminary Literature Review. List at least 4 bibliographic references, cited using APA style. Slide 3: Prepare a knowledge problem components table. Slide 4: Define the dataset variables according to your topic and literature review. Slide 5: Present the confusion matrix components (TP, TN, FP, FN), precision, recall, and accuracy. Explain the meaning of each metric in the context of your research topic.
4. Task 2 (30%). Classification Model Selection.
5. Apply alternative representations of your dependent variable. Normalize or standardize independent variables. Modify model parameters and document each run. Try different classification algorithms and parameter settings. Run at least 15 different experiments and complete the result summary table. Choose the best-performing model(s) and justify your selection. Include the output from the best run in the presentation.
6. Task 3 (20%). Clustering and Final Presentation.
7. Use the k-means algorithm with 3, 4, and 5 clusters. Exclude the target variable during clustering. Characterize each cluster in your own words and in the context of the research topic. Use "Visualize Cluster Assignment" and choose the best scatterplots to illustrate findings. Add explanations and labels for each cluster. Integrate the newly created cluster labels as a new feature in the best classification model. Report on the new evaluation metrics after integration.

## References

Allsopp, D. H., Minskoff, E. H., & Bolt, L. (2005). Individualized course–specific strategy instruction for college students with learning disabilities and ADHD: Lessons learned from a model demonstration project. *Learning Disabilities Research & Practice*, *20*(2), 103–118. [CrossRef]

Auster, E. R., & Wylie, K. K. (2006). Creating active learning in the classroom: A systematic approach. *Journal of Management Education*, *30*(2), 333–353. [CrossRef]

Álvarez-Godos, M., Ferreira, C., & Vieira, M.-J. (2023). A systematic review of actions aimed at university students with ADHD. *Frontiers in Psychology*, *14*, 1216692. [CrossRef] [PubMed]

Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2018). Design of statistics learning environments. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 473–502). Springer International Publishing. [CrossRef]

Binder, T., Sandmann, A., Sures, B., Friege, G., Theyssen, H., & Schmiemann, P. (2019). Assessing prior knowledge types as predictors of academic achievement in the introductory phase of biology and physics study programmes using logistic regression. *International Journal of STEM Education*, *6*(1), 33. [CrossRef]

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, *26*(3–4), 369–398. [CrossRef]

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. [CrossRef]

De Veaux, R. D., Agarwal, M., Averett, M., Baumer, B. S., Bray, A., Bressoud, T. C., Bryant, L., Cheng, L. Z., Francis, A., Gould, R., Kim, A. Y., Kretchmar, M., Lu, Q., Moskol, A., Nolan, D., Pelayo, R., Raleigh, S., Sethi, R. J., Sondjaja, M., . . . Ye, P. (2017). Curriculum guidelines for undergraduate programs in data science. *Annual Review of Statistics and Its Application*, *4*(1), 15–30. [CrossRef]

DuPaul, G. J., Weyandt, L. L., O'Dell, S. M., & Varejao, M. (2009). College students with ADHD: Current status and future directions. *Journal of Attention Disorders*, *13*(3), 234–250. [CrossRef] [PubMed]

Eliyasni, R., Kenedi, A. K., & Sayer, I. M. (2019). Blended learning and project based learning: The method to improve students' higher order thinking skill (HOTS). *Jurnal Iqra': Kajian Ilmu Pendidikan*, *4*(2), 231–248. [CrossRef]

Emmers, E., Jansen, D., Petry, K., van der Oord, S., & Baeyens, D. (2017). Functioning and participation of students with ADHD in higher education according to the ICF-framework. *Journal of Further and Higher Education*, *41*(4), 435–447. [CrossRef]

Fernandes, S. R. G. (2014). Preparing graduates for professional practice: Findings from a case study of project-based learning (PBL). *Procedia—Social and Behavioral Sciences*, *139*, 219–226. [CrossRef]

Frazier, T. W., Youngstrom, E. A., Glutting, J. J., & Watkins, M. W. (2007). ADHD and achievement: Meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students. *Journal of Learning Disabilities*, *40*(1), 49–65. [CrossRef] [PubMed]

Garfield, J., & Ben-Zvi, D. (2009). Helping students develop statistical reasoning: Implementing a statistical reasoning learning environment. *Teaching Statistics*, *31*(3), 72–77. [CrossRef]

Gould, R. (2021). Toward data-scientific thinking. *Teaching Statistics*, *43*(S1), S11–S22. [CrossRef]

Hartung, C. M., Canu, W. H., Serrano, J. W., Vasko, J. M., Stevens, A. E., Abu-Ramadan, T. M., Bodalski, E. A., Neger, E. N., Bridges, R. M., Gleason, L. L., Anzalone, C., & Flory, K. (2022). A new organizational and study skills intervention for college students with ADHD. *Cognitive and Behavioral Practice*, *29*(2), 411–424. [CrossRef]

Karaçalli, S., & Korur, F. (2014). The effects of project-based learning on students' academic achievement, attitude, and retention of knowledge: The subject of "electricity in our lives". *School Science and Mathematics*, *114*(5), 224–235. [CrossRef]

Khalemsky, A., Gelbard, R., & Stukalin, Y. (2024). Constructing a course on classification methods for undergraduate non-STEM students: Striving to reach knowledge discovery. *Journal of Statistics and Data Science Education*, *31*(1), 68–76. [CrossRef]

Khalemsky, A., & Stukalin, Y. (2024). Effect of students' ethnic and cultural backgrounds on the success of practical projects within statistics and data science courses. *Statistics Education Research Journal*, *23*(2), 2. [CrossRef]

Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving Schools*, *19*(3), 267–277. [CrossRef]

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, *35*(6), 382–385. [CrossRef]

MacGillivray, H., & Pereira-Mendoza, L. (2011). Teaching statistical thinking through investigative projects. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 109–120). Springer Netherlands. [CrossRef]

Marks, A., & AL-Ali, M. (2022). Digital transformation in higher education: A framework for maturity assessment. In M. Alaali (Ed.), *COVID-19 challenges to university information technology governance* (pp. 61–81). Springer International Publishing. [CrossRef]

McLaren, B. M., Jiao, P., Ouyang, F., & Alavi, A. H. (Eds.). (2022). *Artificial intelligence in STEM education: The paradigmatic shifts in research, education, and technology*. CRC Press. [CrossRef]

Paechter, M., Macher, D., Martskvishvili, K., Wimmer, S., & Papousek, I. (2017). Mathematics anxiety and statistics anxiety. Shared but also unshared components and antagonistic contributions to performance in statistics. *Frontiers in Psychology*, *8*, 1196. [CrossRef] [PubMed]

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, *49*, 345–375. [CrossRef] [PubMed]

Pilot, Z., Suprise, M., Dinius, C., Olechowski, A., & Habib, R. (2023). Structured peer mentoring improves academic outcomes and complements project-based learning in an introductory research methods and statistics course. *Scholarship of Teaching and Learning in Psychology*, *9*(2), 185–195. [CrossRef]

Ryan, R. M., & Deci, E. L. (2022). Self-determination theory. In *Encyclopedia of quality of life and well-being research* (pp. 1–7). Springer. [CrossRef]

Saenab, S., Yunus, S. R., Saleh, A. R., Virninda, A. N., L, H., & Sofyan, N. A. (2018). Project-based learning as the atmoshphere for promoting students' communication skills. *Journal of Physics: Conference Series*, *1028*(1), 012026. [CrossRef]

Schneider, M., & Simonsmeier, B. A. (2025). How does prior knowledge affect learning? A review of 16 mechanisms and a framework for future research. *Learning and Individual Differences*, *122*, 102744. [CrossRef]

Weyandt, L. L., & DuPaul, G. (2006). ADHD in college students. *Journal of Attention Disorders*, *10*(1), 9–19. [CrossRef] [PubMed]

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. [CrossRef] [PubMed]

Wing-yi Cheng, R., Lam, S., & Chung-yan Chan, J. (2008). When high achievers and low achievers work in the same group: The roles of group heterogeneity and processes in project-based learning. *British Journal of Educational Psychology*, *78*(2), 205–221. [CrossRef] [PubMed]



https://doi.org/10.3390/educsci16091503