

CHAPTER III

RESEARCH DESIGN

3.1 Research Design

This study aims to analyze the perceptions of vocational high school (SMK) students regarding the use of ChatGPT feedback in narrative text development. To achieve this objective, the study will employ a quantitative descriptive survey design. The descriptive survey method is chosen because it enables the systematic collection of data from a large number of respondents, provides standardized measurements, and allows for objective statistical analysis (Creswell & Creswell, 2018).

According to Ary et al. (2010), survey research is a quantitative method used to obtain information that describes existing phenomena by asking individuals about their attitudes, beliefs, opinions, or behaviors. This design enables researchers to generalize findings from a sample to a population systematically. Similarly, Creswell and Creswell (2018) emphasize that descriptive survey research allows for the collection of standardized data to describe opinions or characteristics of a group objectively. The standardization of questions ensures that all respondents receive the same stimulus, thus minimizing bias and enhancing reliability (Bryman, 2015).

The survey design in this study aligns with the research objectives, which focus on quantitatively measuring students' perceptions of ChatGPT feedback in narrative writing. The theoretical foundation of this research is based on the Extended Technology Acceptance Model (TAM) proposed by Afzaal et al. (2025), which integrates Intrinsic Motivation (IM) into the traditional TAM framework. The model consists of four dimensions: Intrinsic Motivation (IM), Perceived Ease of Use (PEOU), Perceived Usefulness (PU), and Behavioral Intention (BI). These constructs explain how students' internal motivation, perceptions of usefulness, and ease of use influence their intention to adopt ChatGPT as a learning support tool.

By employing this framework, the study aims to examine how students perceive the usefulness and ease of using ChatGPT feedback, how their attitudes toward the tool influence their engagement, and whether they intend to continue using it in their writing practices. With this approach, the research is expected to provide empirical insights that can serve as a reference for optimizing the use of artificial intelligence-based tools, particularly ChatGPT, in English language learning at the vocational high school level.

3.2 Population & Sample

The population in this study will consist of all students enrolled in vocational high schools (SMK) in Gresik who have studied narrative text in English class (Creswell & Creswell, 2018). The population includes students across grade levels (X, XI, and XII) who have been exposed to narrative text materials as part of the English curriculum. According to the Ministry of Education, Culture, Research, and Technology (Kemendikdasmen, 2025), there are 63 Vocational High Schools in Gresik with a total of approximately 23,756 students. This population represents potential participants who have

the background knowledge necessary for engaging in narrative writing activities using ChatGPT feedback.

The sample for this study will be determined using purposive sampling, as only students meeting specific and relevant criteria will be considered appropriate for the research focus. The inclusion criteria are designed to ensure that participants possess sufficient experience and exposure in using ChatGPT for narrative writing development. **(1)** The selected vocational high school must have officially implemented the use of Artificial Intelligence (AI) tools—particularly ChatGPT—in English learning activities for at least two consecutive academic years (2023–2025), including classroom-based writing practices or project-based tasks. **(2)** The participating students must have developed at least one narrative text or classroom project using ChatGPT during the current or previous academic year, such as for idea generation, vocabulary enrichment, or grammar checking. **(3)** They must have regular access to personal or school-provided digital devices (smartphones, tablets, or laptops) and a stable internet connection to enable effective use of ChatGPT. **(4)** The students must also demonstrate prior familiarity with ChatGPT as a writing support tool, verified through a pre-screening questionnaire. **(5)** Finally, they must voluntarily agree to participate in the study and provide informed consent to ensure ethical compliance. This purposive sampling approach ensures that data are collected from participants who have sufficient readiness and experience in using ChatGPT for narrative writing, thereby improving the validity and reliability of the study’s descriptive findings.

To determine the minimum sample size, Cochran’s formula for estimating a population proportion with finite population correction was applied, using a 95% confidence level ($Z = 1.96$), a conservative proportion estimate ($p = 0.5$), and a 5% margin of error ($e = 0.05$). The formula is as follows:

$$n_0 = \frac{Z^2 \times p(1-p)}{e^2}$$

Substituting the values:

$$n_0 = \frac{1.96^2 \times 0.5(1-0.5)}{0.05^2} = \frac{3.8416 \times 0.25}{0.0025} = 384.16$$

Because the total population ($N = 15,639$) is finite, the **finite population correction (FPC)** was applied:

$$n = \frac{n_0}{1 + \frac{(n_0 - 1)}{N}}$$

Substituting the values:

$$n = \frac{384.16}{1 + \frac{384.16 - 1}{15,639}}$$

$$n = \frac{384.16}{1 + \frac{383.16}{15,639}} = \frac{384.16}{1 + 0.02449} = \frac{384.16}{1.02449} = 375.12$$

Hence, the minimum sample size required is **375** respondents.

To account for potential non-response and to ensure proportional distribution among participating schools, a 5% buffer was added:

$$n_{final} = 375.12 \times 1.05 = 393.88$$

Finally, after *practical rounding*, the total sample size was set to 395 respondents.

Based on Cochran's formula with finite population correction and a 5% non-response buffer, the final target sample size for this study is 395 vocational high school students out of a total population of 15,639 students from 17 participating schools. This number is considered adequate to ensure representativeness and reliability in measuring students' perceptions of ChatGPT feedback in narrative writing.

By clearly distinguishing the population (all students from the 17 permitted vocational schools in Gresik) from the sample (students with prior ChatGPT writing experience), this study ensures that its findings reflect the authentic perceptions of students who have directly interacted with AI-supported feedback in their English writing learning process.

Table 3.1. Sample Size Determination Using Cochran's Formula

Step/Adjustment	Formula/Calculation	Result
Initial Sample	$n_0 = \frac{Z^2 \times p(1-p)}{e^2} = \frac{1.96^2 \times 0.5(1-0.5)}{0.05^2}$	384.16
Finite Population Correction	$n = \frac{n_0}{1 + \frac{(n_0-1)}{N}} = \frac{384.16}{1 + \frac{384.16-1}{15,639}}$	375.12
Non-Response Adjustment (5%)	$n_{adj} = n \times 1.05 = 375.12 \times 1.05$	393.88
Practical Rounded Sample	-	395

Source: Adapted from Cochran (1977)

3.3 Research Instrument

3.3.1 Questionnaire Design

The primary instrument employed in this study was a questionnaire designed to collect quantitative data in accordance with the survey research approach. The use of questionnaires is widely applied in educational research as it allows researchers to systematically measure perceptions and attitudes across large groups of respondents (Cohen et al., 2018; Creswell & Creswell, 2018).

The questionnaire in this study was adapted and further developed from Lam and Le (2024) and Afzaal et al. (2025), with necessary modifications to suit the research context—students' perceptions of ChatGPT feedback in developing their narrative writing skills. The instrument was constructed based on the Extended Technology Acceptance Model (TAM), which explains users' acceptance of technology through four interrelated dimensions: Intrinsic Motivation (IM), Perceived Ease of Use (PEOU), Perceived Usefulness (PU), and Behavioral Intention (BI).

The questionnaire consists of thirty items divided into two main sections. The first section includes twenty-five close-ended statements measured on a five-point Likert scale, with the following response categories: 1 = Strongly Disagree, 2 =

Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree. These items were designed to assess students' perceptions across the four Extended TAM dimensions.

The construction of this questionnaire is grounded in the Extended Technology Acceptance Model (Afzaal et al., 2025) and recent applications in educational technology research (Lam & Le, 2024; Teo, 2011). In the context of this study, the Intrinsic Motivation dimension measures students' internal drive to use ChatGPT because of curiosity, interest, or enjoyment in learning. The Perceived Usefulness dimension assesses the extent to which students believe that ChatGPT feedback helps improve their narrative writing—particularly in grammar, vocabulary, structure, and coherence. The Perceived Ease of Use dimension captures how easily students can operate and understand ChatGPT's feedback features during writing activities. Lastly, the Behavioral Intention dimension identifies students' willingness to continue using ChatGPT feedback in future writing tasks or to recommend it to others.

To provide a clear overview of how these dimensions and indicators are represented in the questionnaire, the blueprint below summarizes the structure and item distribution developed based on the TAM framework and the context of ChatGPT feedback in narrative writing.

Table 3.2 Blueprint of Students' Perception Questionnaire on ChatGPT Feedback

Dimension	Indicators	Item Numbers	Type of Questionnaire
Intrinsic Motivation (IM)	Students' internal drive to use ChatGPT due to enjoyment, curiosity, and interest in improving their writing skills through feedback.	1-3	Close-ended
Perceived Ease of Use (PEOU)	Ease of understanding, accessing, and applying ChatGPT feedback; clear interface and instructions that make it simple to use.	4-6	Close-ended
Perceived Usefulness (PU)	The benefits of ChatGPT in improving narrative writing performance and understanding of story structure, including orientation (characters, setting, opening), complication (sequencing events, building tension/conflict), resolution (ending), re-orientation (moral/lesson), and language features (grammar, vocabulary, sentence structure, and style).	7-21	Close-ended

Behavioral Intention (BI)	Writing actions and intentions based on ChatGPT feedback, such as revising stories, comparing feedback, and plans to use ChatGPT in future writing tasks.	22–25	Close-ended
---------------------------	---	-------	-------------

Note. The questionnaire blueprint was adapted from Lam and Le (2024) and further aligned with the Extended Technology Acceptance Model (TAM) by Afzaal et al. (2025), focusing on students’ perceptions of ChatGPT feedback in narrative writing across four constructs: Intrinsic Motivation (IM), Perceived Ease of Use (PEOU), Perceived Usefulness (PU), and Behavioral Intention (BI).

This instrument design provides a comprehensive understanding of how students perceive ChatGPT’s role in their narrative writing process, focusing not only on its usefulness and ease of operation but also on students’ attitudes and intentions to adopt it further.

3.3.2 Validity and Reliability of the Instrument

3.3.2.1 Validity of the Instrument

This study employed content validity to ensure that the questionnaire accurately represented the constructs being measured—students’ perceptions of ChatGPT feedback in narrative writing based on the Extended Technology Acceptance Model (TAM). According to Cohen et al. (2018), content validity refers to the extent to which the items of an instrument adequately cover the domain of the concept it aims to measure.

The process of establishing content validity was conducted through expert judgment involving two experts: (1) an English teacher from a vocational high school and (2) a university lecturer specializing in writing. In the first stage, the English teacher reviewed the questionnaire for clarity, language appropriateness, and the suitability of the statements for vocational students. Based on the teacher’s feedback, several linguistic and contextual adjustments were made to ensure that each item was easily understood by students.

In the second stage, the revised version was evaluated by the university lecturer to verify its alignment with the four dimensions of the Extended TAM—Intrinsic Motivation (IM), Perceived Ease of Use (PEOU), Perceived Usefulness (PU), and Behavioral Intention (BI). The lecturer assessed whether each item corresponded logically and conceptually to its intended dimension and whether the overall construct validity was consistent with the Extended TAM framework (Afzaal et al., 2025). Since all items were deemed relevant and theoretically appropriate, no further revisions were required.

This approach was considered sufficient because expert judgment is widely accepted for establishing content validity of educational instruments, ensuring that all items accurately represent the theoretical constructs (LAWSHE, 1975). No statistical tests were conducted, as the focus of this

validation was solely on the theoretical and conceptual representativeness of the items.

Through this two-stage validation process, the questionnaire was confirmed to have strong content validity, as each item was judged to be representative, clear, and aligned with the theoretical dimensions of the Extended Technology Acceptance Model in the context of ChatGPT feedback for narrative writing.

3.3.2.2 Reliability of the Instrument

The reliability of the questionnaire was examined through a pilot test conducted at SMKN 1 Dudusampeyan Gresik, involving 20 students from Class XI AK2, who possessed similar characteristics to the target population of the main study. The primary purpose of this pilot test was to assess the internal consistency of the questionnaire items and to ensure that the instrument generated stable and dependable responses.

The reliability coefficient was computed using Cronbach's Alpha with the assistance of SPSS software. Only the close-ended Likert-scale items (Items 1–25) were included in the reliability analysis.

The SPSS output produced an overall Cronbach's Alpha coefficient of 0.846, indicating a high level of internal consistency among the questionnaire items. According to Creswell and Creswell (Creswell & Creswell, 2018), a coefficient value of 0.70 or higher is generally considered acceptable for research instruments. Therefore, the questionnaire used in this study can be regarded as reliable for data collection in the main study. Detailed item-by-item reliability results are presented in the Appendix.

Table 3.3 Reliability Statistics of the Questionnaire

Reliability Statistics	
Cronbach's Alpha	N of Items
.846	25

Source: SPSS Output, October 2025.

3.4 Data Collection

Data collection in this study was conducted using questionnaires as the main instrument, in accordance with the quantitative survey research design. The data collection process began with coordination with the school where the research was conducted to obtain official permission to conduct the research and to ensure the availability of respondents who met the inclusion criteria. After obtaining permission, the researchers socialized the research objectives, questionnaire filling procedures, and explained data confidentiality and informed consent to the students.

The questionnaire was then distributed directly to respondents who met the criteria, namely vocational high school students in Gresik who were allowed to use ChatGPT in their learning process and had experience creating or developing narrative texts with the help of ChatGPT. Data collection was carried out simultaneously in each school to

maintain consistency in the time of completion and minimize the influence of external variables. Each respondent was given sufficient time to complete the questionnaire independently without pressure so that the answers provided reflected their true perceptions.

After all questionnaires were collected, the researchers checked the completeness of the answers to ensure that no data was missing or left blank. Valid data was then coded and entered into data processing software for statistical analysis. This data collection procedure followed the principles of quantitative data collection recommended by Creswell (2018) and Sugiyono (2015), namely systematic, objective, and standardized so that the research results could be trusted and generalized.

3.5 Data Analysis

The data obtained from the close-ended questionnaire were analyzed using both descriptive and inferential statistical techniques. Each item was coded according to the five-point Likert scale, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). The collected data were coded and subsequently processed using SPSS.

3.5.1 Descriptive Statistical Analysis

To answer Research Question 1, descriptive statistics were employed to calculate the frequency, percentage, mean score, and standard deviation of each item and each dimension of the Extended Technology Acceptance Model (TAM), namely Intrinsic Motivation (IM), Perceived Ease of Use (PEOU), Perceived Usefulness (PU), and Behavioral Intention (BI).

According to Pimentel (2010), Likert-scale data with five response categories can be analyzed using mean scores to describe respondents' overall perceptions. Therefore, the mean scores obtained in this study were interpreted using the classification presented in Table.

Table 3.4 Mean Score Interpretation of Five-Point Likert Scale

Mean Score Range	Interpretation Category
1.00 – 1.80	Strongly Disagree
1.81 – 2.60	Disagree
2.61 – 3.40	Neutral/Uncertain
3.41 – 4.20	Agree
4.21 – 5.00	Strongly Agree

Source: (Pimentel, 2010)

This categorization was used to determine the overall level of students' perceptions toward the use of ChatGPT in supporting narrative writing development.

Through descriptive analysis, this study aimed to illustrate general trends in students' responses, including their motivation to use ChatGPT, their perceived ease of interacting with the tool, their perceived usefulness of ChatGPT for narrative writing, and their intention to continue using ChatGPT in the future.

3.5.2 Inferential Statistical Analysis

To answer Research Question 2, an inferential statistical test, specifically the Pearson Product-Moment Correlation, was employed. This statistical test was used

to examine the relationships among Intrinsic Motivation (IM), Perceived Ease of Use (PEOU), Perceived Usefulness (PU), and Behavioral Intention (BI).

According to Pallant (2020), Pearson correlation analysis is appropriate for examining the strength and direction of relationships between continuous variables. The strength of the correlation coefficients in this study was interpreted based on the criteria shown in Table.

Table 3.5 Interpretation of Pearson Correlation Coefficient Values

Correlation Coefficient (r)	Interpretation
0.00 – 0.19	Very weak
0.20 – 0.39	Weak
0.40 – 0.59	Moderate
0.60 – 0.79	Strong
0.80 – 1.00	Very strong

Source: Pallant, (2020)

This analysis allowed the researcher to identify the strength and direction of the relationships among the four TAM dimensions and to determine which variables had the strongest association with students' behavioral intention to use ChatGPT.

The combined use of descriptive and inferential statistical analyses ensured a comprehensive understanding of students' perceptions and the interrelationships among the dimensions of the acceptance model in the context of ChatGPT-assisted narrative writing development.

