



Enhancing Conceptual Coherence: Students' Mental Models of the Male Reproductive System through *Collaborative Constructive* Approach

Anindya Utami Putri¹, Adi Rahmat¹, Yanti Hamdiyati¹

¹Biology Education, Faculty of Mathematics and Natural Sciences Education (FPMIPA), Universitas Pendidikan Indonesia

*Corresponding Author: Adi Rahmat

E-mail: adirahmat@upi.edu



Article Info

Article history:

Received 30 May 2026

Received in revised form 15

June 2026

Accepted 20 August 2026

Keywords:

Collaborative

Constructive

Mental Model

Male Reproductive System

Collaborative Learning

Abstract

This study aims to analyze the effect of the Collaborative Constructive approach on the quality of students' mental models in the topic of the male reproductive system. The study employed a mixed-methods approach with an embedded experimental design and a quasi-experimental pretest-posttest control group design. The research sample consisted of 73 eleventh-grade students at a public senior high school in Cirebon City, with 37 students assigned to the experimental class and 36 students to the control class. The experimental class participated in Collaborative Constructive learning through worksheet exploration, group discussion, completion of incomplete concept maps, presentations, and concept clarification, whereas the control class received teacher-centered instruction. The research instrument was an open-ended essay test assessing the aspects of content, prediction, explanation, and drawing. The data were analyzed using normalized gain scores, the Mann-Whitney test, and qualitative analysis of changes in students' responses. The results showed that the experimental class obtained a normalized gain score of 0.74, which falls into the high category, while the control class obtained a score of 0.51, categorized as moderate. The Mann-Whitney test indicated a significant difference in the improvement of mental models between the two classes. Qualitatively, the experimental class demonstrated a more directed shift in mental models toward scientific understanding, particularly in the aspects of content, prediction, and explanation, although the visual representation aspect still requires further reinforcement. These findings indicate that the Collaborative Constructive approach can support the reconstruction of students' mental models in abstract and complex biological topics.

Introduction

A meaningful understanding of biological concepts is not reflected merely in students' ability to recall factual information, but also in their capacity to construct mental models that allow them to explain biological phenomena scientifically (Johnson et al., 2026; Sanchez et al., 2026; Vergara et al., 2026). Mental models serve as cognitive representations through which students interpret, predict, and explain phenomena based on their existing conceptual knowledge (Johnson-Laird, 1983). Since mental models are internal constructs, students' understanding must be elicited through external representations, such as written responses, drawings, or concept maps (Hamdiyati et al., 2018). This is particularly important in biology learning, where many concepts require students to coherently connect biological structures, functions, processes, and mechanisms.

One topic that demands such an integrated understanding is the human reproductive system, particularly the male reproductive system (Stavros et al., 2025; Antinozzi et al., 2025). This

topic encompasses the anatomy of reproductive organs, sperm transport pathways, spermatogenesis, and reproductive disorders, most of which are not directly observable (Lata & Hajam, 2026). Its abstract and complex nature makes students susceptible to fragmented understanding, misconceptions, or even a Blank Mental Model (BMM), a condition in which students are unable to develop meaningful conceptual representations. Wahyuni et al. (2019) reported that students often struggle to understand the human reproductive system due to the large number of scientific terms, the abstract nature of the concepts, monotonous instruction, and limited access to learning resources. Similar findings were reported by Salsabilla and Melta (2024), who identified misconceptions among senior high school students across several subconcepts of human reproduction.

The limited quality of students' mental models in learning the reproductive system represents an important pedagogical issue, as students are expected not only to recognize organ names or biological terminology, but also to understand the relationships among structure, function, process, and the consequences of reproductive disorders in a scientific manner (Lin et al., 2025; Zhang et al., 2025). Incomplete mental models may hinder students' ability to explain biological mechanisms coherently (Seel, 2017). However, previous studies have largely focused on measuring learning outcomes, examining learning difficulties, or identifying misconceptions, while relatively few have specifically investigated changes in students' mental models following a learning intervention. Moreover, studies that integrate diagnostic mental model mapping with instructional strategies designed to reconstruct students' understanding of the male reproductive system at the senior high school level remain limited (Baharuddin & Umar, 2026; Dong et al., 2026; Neway, 2026).

One instructional approach that may address this issue is Collaborative Constructive learning. Rooted in social constructivism, this approach positions student interaction as a medium for constructing, testing, and revising understanding through discussion, argumentation, and scaffolding within the Zone of Proximal Development (ZPD) (Eun, 2017). In this study, Collaborative Constructive learning was implemented through student worksheet exploration, group discussions, the completion of partially developed concept maps, and presentations of group discussion results (Shetty et al., 2025; Sholihah et al., 2025; Chiu & Hwang, 2026). These activities provided students with opportunities to organize their knowledge, identify conceptual inaccuracies, and communicate their representations of thinking more explicitly. The use of concept maps further supported students in recognizing relationships among concepts, enabling their understanding to move beyond factual memorization toward more structured mental models (Novak & Gowin, 1984).

Based on these research gaps, this study aims to analyze the effect of implementing Collaborative Constructive learning on the quality of eleventh-grade students' mental models in the topic of the male reproductive system. The novelty of this study lies in its use of mental models as the primary indicator of learning success through the diagnostic framework of Content, Prediction, Explanation, and Drawing (C-P-E-D), adapted from Kurnaz and Eksi (2015). This framework enables students' understanding to be mapped not only in terms of content mastery, but also in relation to their ability to predict, explain, and visually represent biological concepts. Accordingly, this study hypothesizes that students who participate in Collaborative Constructive learning will demonstrate greater improvement in the quality of their mental models than those who receive teacher-centered instruction (Sunarto, 2025; Rachman et al., 2026; Lin & Hung, 2026).

Methods

Research Design

This study employed a mixed-methods approach using an embedded experimental design. This design was adopted because the primary aim of the study was to test a causal effect of the Collaborative Constructive approach, which required quantitative comparison as the main framework, while qualitative data were embedded within this framework to explain the mechanisms underlying the quantitative findings rather than to stand as an independent line of inquiry. Quantitative data were used to examine differences in the improvement of students' mental models between the experimental and control classes, while qualitative data were used to describe patterns of change in students' responses based on mental model categories. Specifically, the qualitative data were derived from the same open-ended instrument used for the quantitative analysis, in the form of students' response patterns and their shifts across mental model levels at the pretest, practice, and posttest stages. These qualitative patterns were used to illustrate which specific aspects content, prediction, explanation, or drawing accounted for the numerical differences identified through the N-Gain and Mann-Whitney analyses, so that the qualitative findings served to support and contextualize, rather than parallel, the quantitative results. Quantitatively, this study adopted a quasi-experimental design with a pretest-posttest control group design. This design was selected because the study was conducted in pre-existing classes, making individual randomization impractical.

Population and Sample

The study was conducted among eleventh-grade students at a public senior high school in Cirebon City. The sample was selected using purposive sampling by considering similarities in grade level, curriculum, learning material, and the willingness of the school and teachers to participate in the study. Class selection was based on the recommendation of the biology teacher, who considered class characteristics, students' prior daily test results, and the suitability of the class's academic track with a biology-focused learning emphasis. Following this recommendation, Class XI C2 was designated as the experimental class and Class XI C1 as the control class. The sample consisted of 73 students, comprising 37 students in the experimental class and 36 students in the control class. All students enrolled in both classes participated from the beginning to the end of the study; no students were excluded or absent during the pretest, learning intervention, or posttest, so that the number of participants remained unchanged throughout. The equivalence of initial conditions between the two classes was considered based on the teacher's recommendation, students' prior daily test results, and their pretest scores, so that differences in the final outcomes could be more confidently attributed to the learning intervention rather than to pre-existing differences in students' initial ability.

Learning Process

Both classes were taught by the same biology teacher, so that differences in learning outcomes could be attributed to the instructional approach rather than to differences between teachers. The two classes received the same learning material and were allocated the same amount of instructional time, so that the only difference between them was the instructional approach applied. The experimental class received instruction using the Collaborative Constructive approach, whereas the control class received teacher-centered instruction. The learning materials covered three subtopics of the male reproductive system, namely the anatomy of reproductive organs, sperm transport pathways, and disorders of the reproductive system. The learning process was carried out over three meetings conducted in April 2026, with each meeting lasting approximately two lesson periods followed by one additional lesson period,

adjusted to the school's scheduling. In the experimental class, learning activities were carried out through student worksheet exploration, group discussions, the completion of partially developed concept maps, presentations of discussion results, and concept clarification by the teacher. In the experimental class, the worksheet (LKPD) consisted of five open-ended essay items; several items included sub-questions (e.g., parts a and b) designed to help students develop their answers progressively. Students worked in groups to answer the worksheet items and subsequently completed a partially developed concept map, in which some parts of the concept relationships were already provided as scaffolding to help students complete the remaining sections. Each group then presented their discussion results, followed by a question-and-answer session and responses from other students. At the end of each meeting, the teacher provided concept clarification by correcting misconceptions, completing incomplete answers, and displaying a PowerPoint presentation to reinforce the concepts discussed during the presentation session. Throughout this process, the teacher acted as a facilitator and co-researcher, while the researcher was involved in implementing the lesson, observing the learning process, collecting data, and documenting research activities. In the control class, learning was conducted through teacher explanation, question-and-answer sessions, note-taking, and practice exercises, without the use of partially developed concept maps or structured collaborative discussion.

Research Instrument

The research instrument was an open-ended essay test developed based on the mental model aspects of Content, Prediction, Explanation, and Drawing (C-P-E-D), adapted from Kurnaz and Eksi (2015). The Content aspect was used to measure students' mastery of factual concepts, Prediction to assess their ability to anticipate the effects of a biological condition, Explanation to evaluate the scientific reasoning underlying their predictions, and Drawing to assess their ability to visually represent concepts. Each aspect (C, P, E, D) was scored using a three-point rubric: a score of 0 was given when a student left the item blank; a score of 1 was given for an incorrect or incomplete response (e.g., an inaccurate prediction, a weak explanation that merely restated the question, or a drawing containing conceptual errors); and a score of 2 was given for a fully correct response reflecting accurate scientific understanding. The initial instrument consisted of six sets of open-ended essay questions. Each set contained four question components based on the C-P-E-D aspects, allowing students' responses to be analyzed in terms of content, prediction, explanation, and visual representation. Instrument validity was tested using Pearson Product-Moment correlation with the assistance of SPSS Statistics 25. The results of the instrument testing indicated that five sets of questions were valid, while one set was revised. In the implementation of the study, three sets of questions were used as the pretest and posttest instruments in accordance with the indicator coverage of the three subtopics. Instrument reliability was tested using Cronbach's alpha and yielded a coefficient of 0.938, indicating that the instrument had very high reliability.

To illustrate how individual responses were translated into mental model levels, a student who obtained a score profile of C = 2, P = 2, E = 1, D = 0 correct content and prediction, but a weak explanation and no drawing was classified as a Predictive-Content Mental Model (PCMM). Profiles in which all four aspects were scored 2 were classified as the Scientifically Accepted Mental Model (SAMM), whereas profiles in which all four aspects were scored 0 were classified as the Blank Mental Model (BMM).

Data Analysis

To formulate the quality of students' understanding, student profiles were represented in a one-dimensional matrix, $MM = [C, P, E, D]$. Each aspect in the matrix was scored as 0 if no response

was provided, 1 for an incorrect or incomplete response, and 2 for a correct response. Based on the combination of these scores, cognitive profiles were classified into six hierarchical levels. Level 1 represented the Scientifically Accepted Mental Model (SAMM), indicating fully scientific understanding across all aspects. Levels 2 to 4 were categorized as Hybrid Mental Models, with different proportions of correct aspects: Level 2 indicated three correct aspects, Level 3 indicated two correct aspects, and Level 4 indicated one correct aspect. Level 5 represented the Unscientifically Accepted Mental Model (USAMM), indicating understanding that was inconsistent with scientific concepts, while Level 6 represented the Blank Mental Model (BMM), assigned when students had no identifiable mental model (Kurnaz & Eksi, 2015).

Data processing began with the scoring and coding of students' responses. The shift in students' profiles from pretest to posttest was then analyzed descriptively to determine the percentage distribution of understanding in each group. Subsequently, students' improvement in each class was quantified using the normalized gain (N-Gain) score, calculated as $N\text{-Gain} = (\text{Posttest score} - \text{Pretest score}) / (\text{Maximum possible score} - \text{Pretest score})$ (Hake, 1998), with resulting scores categorized as high ($N\text{-Gain} > 0.7$), moderate ($0.3 \leq N\text{-Gain} \leq 0.7$), or low ($N\text{-Gain} < 0.3$). Prior to hypothesis testing, the N-Gain scores were tested for normality using the Shapiro-Wilk test and for homogeneity of variance using Levene's test. Because the N-Gain scores were not normally distributed, the non-parametric Mann-Whitney U test was applied to the N-Gain scores rather than to the raw posttest scores to examine whether the improvement in mental model quality differed significantly between the experimental and control classes, using a significance level of $\alpha = 0.05$. The magnitude of this difference was further quantified using the effect size r , calculated as $r = Z / \sqrt{N}$, where Z is the standardized test statistic from the Mann-Whitney U test and N is the total sample size. This quantitative analysis was complemented by qualitative analysis to identify patterns of misconceptions that continued to emerge during the learning process. In addition to pretest and posttest data, students' responses during the practice stage were also analyzed descriptively to examine the process of mental model shifts throughout the learning activities.

Ethical Considerations

This study obtained formal research permits from the university and from the school where the research was conducted. Informal permission was also obtained from the biology teacher, given that formal permission had already been granted by the school. Prior to the administration of the pre-test and post-test, students were given an explanation of the purpose of the activity, the procedures involved, and the use of their data for research purposes; students' agreement to participate was obtained verbally prior to each data collection activity. To protect students' privacy, all student data were anonymized using participant codes, so that no personally identifiable information was disclosed in the article or in the data analysis.

Results and Discussion

This study conducted hypothesis testing to determine the effect of the Collaborative Constructive approach on the improvement of students' mental models in the topic of the male reproductive system. The results of the assumption tests showed that the data were not normally distributed but had homogeneous variances. Descriptively, the experimental class obtained a pretest mean of 4.30 (SD = 3.49, median = 4.00, range = 0–20) and a posttest mean of 19.19 (SD = 4.14, median = 21.00, range = 5–24), while the control class obtained a pretest mean of 3.42 (SD = 3.43, median = 3.00, range = 0–12) and a posttest mean of 13.97 (SD = 5.46, median = 14.50, range = 1–22). The Shapiro-Wilk test indicated that the posttest scores of the experimental class were not normally distributed ($W = .798$, $p < .001$), whereas those of the

control class approximated a normal distribution ($W = .950$, $p = .107$); the N-Gain scores followed the same pattern, being non-normally distributed in the experimental class ($W = .756$, $p < .001$) but normally distributed in the control class ($W = .962$, $p = .250$). Levene's test indicated that the variances of the pretest ($F = 0.98$, $p = .327$), posttest ($F = 3.79$, $p = .056$), and N-Gain scores ($F = 0.27$, $p = .604$) were homogeneous between the two classes. Therefore, the analysis of differences in students' mental model improvement was continued using the non-parametric Mann-Whitney U test.

Table 1. *Mann-Whitney U Test*

Data	Kelas	N	Mean Rank	Sum of Ranks	U	Z	p-value	Interpretation
Mental Model	Experimental	37	47.45	1755.50	279.5	-4.267	$p < 0.001$	There was a significant difference in improvement
	Control	36	26.26	945.50				

The results of the Mann-Whitney U test presented in Table 1 indicate a significant difference in the improvement of mental model quality between the experimental and control classes.

Table 2. N-Gain Test

Kelas	N-Gain	Interpretation
Experimental	0.74	High
Control	0.51	Moderate

Based on Table 2, the N-Gain score of students' mental models in the experimental class was categorized as high, while that of the control class was categorized as moderate

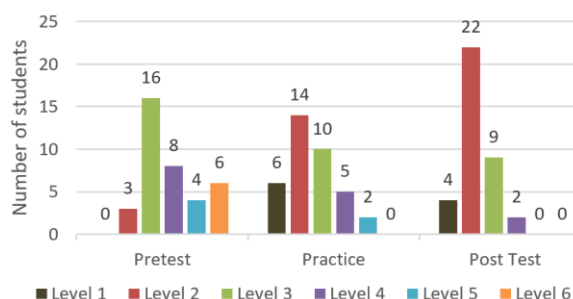


Figure 1. *Mental Models in the Reproductive Organ Subtopic in the Experimental Class*

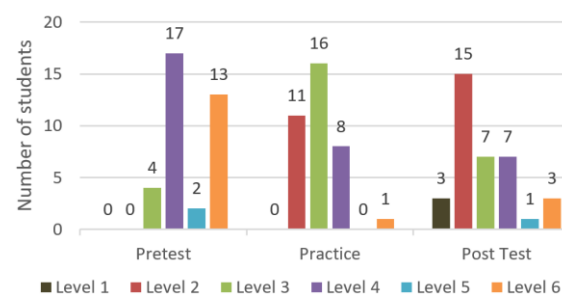


Figure 2. *Mental Models in the Reproductive Organ Subtopic in the Control Class*

Based on Figure 1 and Figure 2, students' initial understanding of the reproductive organ subtopic in both classes was still dominated by partial mental models, particularly at Level 3 and Level 4 (pretest: Level 3 = 48.6% and Level 4 = 21.6% in the experimental class; Level 4 = 52.8% and Level 6 = 30.6% in the control class). During the practice stage, the experimental class began to shift toward more scientific levels of mental models, whereas the control class still showed a more varied distribution of understanding. In the posttest, the experimental class was dominated by Level 2, namely the Non-Drawing Mental Model (NDMM), while several students in the control class remained at lower levels of mental models (posttest: Level 2 = 59.5% in the experimental class vs. 47.2% in the control class; Levels 4–6 combined = 24.8% in the control class vs. 5.4% in the experimental class). This pattern indicates that students in

the experimental class were better able to develop conceptual and causal understanding, although their visual representation of anatomy was not yet fully optimal.

In contrast, the development of mental models in the control class tended to be more limited because teacher-centered instruction placed greater emphasis on the reception of information from the teacher. This condition may lead students to recognize reproductive organs as separate components without being able to fully relate them to their functions and biological consequences. This is consistent with the view that physiological understanding requires the ability to connect structure, process, and function, rather than merely recalling components separately (Lira & Gardner, 2020). The dominance of NDMM in the experimental class also confirms that students' verbal ability to explain concepts was not yet fully accompanied by their ability to depict anatomical structures accurately. Therefore, learning about reproductive organs still needs to be strengthened through visual representation exercises, as students' drawings can reveal gaps in understanding that may not always be evident in verbal responses (Walton, 2023).

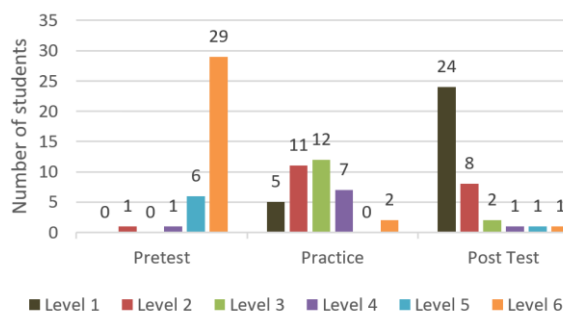


Figure 3. Mental Models in the Sperm Pathway Subtopic in the Experimental Class

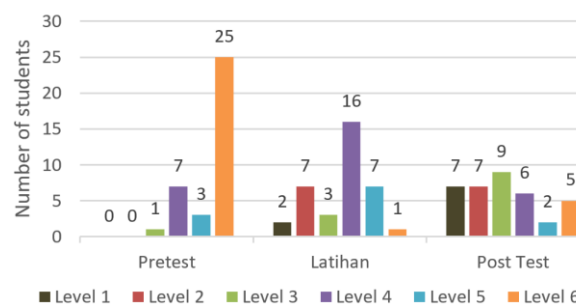


Figure 4. Mental Models in the Sperm Pathway Subtopic in the Control Class

Based on Figure 3 and Figure 4, students' initial understanding of the sperm pathway subtopic in both classes was still dominated by Level 6, or the Blank Mental Model (BMM) (pretest: Level 6 = 78.4% in the experimental class and 69.4% in the control class). This condition indicates that students did not yet have adequate conceptual representations of the sperm pathway as a sequential biological process. During the practice stage, the experimental class began to shift toward the Hybrid Mental Model category, particularly Level 2 and Level 3, whereas the control class remained dominated by partial mental models. In the posttest, the experimental class was dominated by Level 1, or the Scientifically Accepted Mental Model (SAMM), while the control class remained distributed across several levels of mental models.

In contrast, the distribution of mental models in the control class suggests that teacher-centered instruction had not fully helped students integrate structure, process, and function in the sperm pathway. Students may recognize several organs through which sperm passes, but still have difficulty explaining the mechanistic relationships among them, such as why cutting the vas deferens during vasectomy prevents sperm from being released without stopping sperm production in the testes. This type of understanding requires connections among structural components, mechanisms of action, and resulting functions, as explained in the structure-behavior-function framework (Gregoric & Torkar, 2022). In this context, concept maps serve as cognitive tools because they help students organize concepts visually and relationally. The meta-analysis by Anastasiou et al. (2024) also showed that concept maps are effective in improving science learning achievement because they help students arrange information into a more organized knowledge structure. Thus, the findings on the sperm pathway subtopic

indicate that Collaborative Constructive learning more strongly supports the formation of scientific mental models by providing space for concept construction, pathway visualization, and revision of understanding through social interaction.

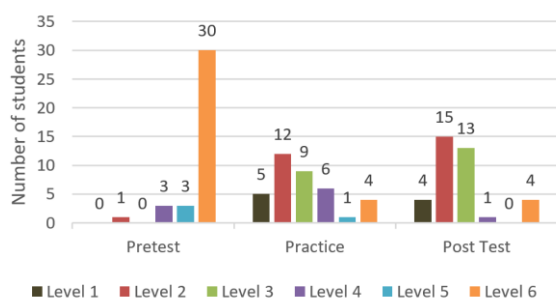


Figure 5. Mental Models in the Reproductive Disorder Subtopic in the Experimental Class

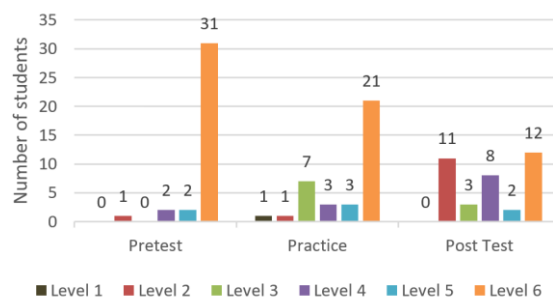


Figure 6. Mental Models in the Reproductive Disorder Subtopic in the Control Class

Based on Figure 5 and Figure 6, students' initial understanding of the reproductive disorder subtopic in both classes was dominated by Level 6, or the Blank Mental Model (BMM) (pretest: Level 6 = 83.8% in the experimental class and 86.1% in the control class). This condition reflects that students did not yet have adequate conceptual representations to explain disorders of the male reproductive system. During the practice stage, the experimental class began to shift toward Level 2, or the Non-Drawing Mental Model (NDMM), and Level 3 as part of the Hybrid Mental Model category, while the control class remained dominated by BMM. In the posttest, the experimental class was still dominated by Level 2 and Level 3 and began to show achievement at Level 1, or the Scientifically Accepted Mental Model (SAMM). In contrast, the control class still showed a dominance of BMM and did not reach Level 1 (posttest: Level 2 = 43.2% and Level 3 = 32.4% in the experimental class vs. Level 2 = 33.3% in the control class; BMM = 10.8% in the experimental class vs. 30.6% in the control class).

The stronger shift in the experimental class may be associated with Collaborative Constructive activities, which provided students with opportunities to discuss cause-and-effect relationships, such as the relationships among infection, inflammation, reproductive tract disorders, and their effects on system function. Group discussions and partially completed concept maps helped students organize these causal relationships in a more directed manner. In contrast, teacher-centered instruction in the control class tended to position students as recipients of separate pieces of information, so their understanding of reproductive disorders had not yet developed into a complete mental model. The findings in the reproductive disorder subtopic show that Collaborative Constructive learning supported the strengthening of the Content, Prediction, and Explanation aspects, although some students still required visual representation exercises to depict the location and mechanisms of disorders more accurately.

These differences indicate that each subtopic places different cognitive demands on students. Reproductive organs require visual representation, the sperm pathway requires an understanding of process sequencing, and reproductive disorders require cause-and-effect reasoning. The dominance of partial mental models, such as NDMM and the Predictive-Content Mental Model (PCMM), indicates that some students were already able to master content and make predictions, but were not yet fully able to construct complete scientific explanations or visual representations. This pattern is consistent with Hamdiyati et al. (2022), who emphasized that students' mental models in biology should be examined through the integration of verbal, visual, and relational representations. Therefore, Collaborative Constructive learning encouraged a shift in students' mental models toward more scientific

understanding, although the Drawing and Explanation aspects still require further strengthening.

Table 3. Representation of the Trajectory of Changes in Students' Responses from Pretest to Posttest

Sub Topic	Class	Pretest	Posttest
Reproductive Organs	Experimental	Content: "Penis: an organ for releasing sperm and involved in the reproductive process. Testis: produces testosterone. Epididymis: the site of sperm maturation and storage. Urethra: the channel through which urine and sperm exit." Prediction: "The number of sperm decreases, and the sperm do not mature properly." Level 3: Predictive-Content Mental Model	Content: "Testis: produces sperm. Urethra: the channel through which sperm exits. Penis: a copulatory organ that delivers sperm into the vagina." Prediction: "Sperm production decreases or its quality becomes poor, resulting in a low sperm count, abnormal shape, and a reduced chance of fertilization." Explanation: "Disorders of the testes impair sperm production because temperature increases, testosterone decreases, and free radicals damage sperm cells." Level 2: Non-Drawing Mental Model
	Control	Content: "Testis: the site of sperm cell formation. Epididymis: a temporary storage site for sperm cells. Vas deferens: connects the epididymis." Level 4: Content-Only Mental Model	Content: "Testis: the site of sperm formation. Epididymis: a temporary storage site for sperm. Vas deferens: a channel that connects the epididymis to the urethra." Prediction: "It will decrease significantly." Explanation: "This mechanism of quality decline occurs because the testes are the main site of sperm production." Level 2: Non-Drawing Mental Model
Sperm Pathway	Experimental	Left blank. (Level 6)	Content: "Testis: the site where sperm are produced. Epididymis: the site where

Sub Topic	Class	Pretest	Posttest
			sperm mature or are stored. Vas deferens: transports sperm from the epididymis to the urethra. Ejaculatory duct: the channel where sperm mixes with fluid from the seminal vesicles. Urethra: the final channel that releases sperm from the body through the penis.” Prediction: “The vas deferens is cut, so sperm from the epididymis cannot pass into the urethra. Sperm continue to be produced in the testes but only reach the epididymis and are then absorbed by the body. During ejaculation, pregnancy does not occur because sperm cannot exit to fertilize the egg.” Explanation: “Vasectomy cuts the vas deferens so that sperm cannot pass into the urethra and are eventually absorbed by the body, while ejaculation still occurs without sperm.” Drawing: The student drew a diagram of the sperm pathway from the testis – epididymis – vas deferens – ejaculatory duct – urethra – out of the body. Level 1: Scientifically Accepted Mental Model
	Control	Left blank. (Level 6)	Content: “Testis: the site of sperm formation. Epididymis: the channel for sperm maturation. Vas deferens: transports sperm to the urethra. Urethra: the final channel that carries sperm out through the penis.” Prediction: “The pathway is cut, sperm cannot exit, and

Sub Topic	Class	Pretest	Posttest
			sperm can be reabsorbed by the body.” Level 3: Predictive-Content Mental Model
Reproductive Disorders	Experimental	Left blank. (Level 6)	Content: “Gonorrhea is a sexually transmitted disease caused by the bacterium <i>Neisseria gonorrhoeae</i> . Its effect is inflammation of the tract.” Prediction: “It will be disrupted or obstructed, with inflammation and swelling in the urethral tract.” Explanation: “Inflammation causes the urethral wall to swell and produce pus.” Level 2: Non-Drawing Mental Model
	Control	Left blank. (Level 6)	Left blank. (Level 6)

This finding suggests that Collaborative Constructive learning contributed more strongly to the improvement of students’ mental models than teacher-centered instruction. This finding reinforces the results of the Mann-Whitney U test, indicating that Collaborative Constructive learning provided stronger support for improving students’ mental models than teacher-centered instruction. The difference in improvement shows that students’ engagement in student worksheet exploration, group discussions, completion of partially developed concept maps, and presentations of discussion results played a role in helping them construct more structured understanding.

This difference may be attributed to the characteristics of Collaborative Constructive learning, which provides opportunities for students to construct and revise their understanding through student worksheet exploration, group discussion, and the completion of partially completed concept maps. In the reproductive organ subtopic, students are not expected merely to memorize the names of organs, but also to understand the relationships among structure, function, and biological mechanisms. Collaborative activities allow students to compare ideas, clarify conceptual errors, and receive scaffolding from both peers and the teacher (Eun, 2017). In addition, concept maps help students explicitly organize relationships among concepts, enabling previously fragmented knowledge to develop into a more coherent structure of understanding (Schroeder et al., 2017).

The stronger shift observed in the experimental class may be related to the nature of Collaborative Constructive learning, which helps students organize the sperm pathway as an interconnected system. Through group discussion and partially completed concept maps, students were guided to understand the sequential relationships among the testis, epididymis,

vas deferens, ejaculatory duct, urethra, and penis, rather than merely memorizing organ names. In this subtopic, students need to understand the processes of sperm formation, maturation, transport, and release as a connected sequence. This is in line with Momsen et al. (2022), who emphasized that biology learning should develop systems thinking so that students can understand the relationships among components within a biological system.

This pattern indicates that the reproductive disorder subtopic requires more complex reasoning than the identification of reproductive organs or the sperm pathway. Students must understand the relationships among the normal condition of organs, causes of disorders, changes in function, and their effects on the reproductive system. This difficulty is consistent with the findings of Nasution and Khairuna (2022), who showed that the highest level of misconception in the topic of the human reproductive system occurred in the reproductive disorder subtopic. Andariana et al. (2020) also found that misconceptions in human anatomy and physiology frequently appear in the reproductive system subtopic, indicating the need for learning that can more explicitly connect the concepts of structure, function, and mechanism.

Comparatively, the shifts in students' mental models showed different patterns across subtopics. In the reproductive organ subtopic, many students in the experimental class remained in the Non-Drawing Mental Model (NDMM) category, confirming that students' verbal understanding was not yet fully accompanied by their ability to represent anatomy visually. In the sperm pathway subtopic, the shift in mental models appeared to be the strongest because students were supported by activities that required them to arrange sequences and relationships among organs using partially completed concept maps. In contrast, reproductive disorders remained the most problematic subtopic because the Blank Mental Model (BMM) still appeared, particularly in the control class.

Based on Table 3, differences in the quality of students' mental models were evident not only in changes in level, but also in the depth of students' responses. In the reproductive organ subtopic, both the experimental and control classes reached Level 2 in the posttest, but the quality of their explanations differed. The responses of students in the experimental class showed more specific causal relationships among organ disorders, physiological changes, sperm quality, and the chance of fertilization. In contrast, students' responses in the control class remained general because they only linked testicular function to decreased sperm quality. This shows that the same mental model category does not necessarily represent the same depth of reasoning. In the context of physiology, strong understanding requires connections among structure, mechanism, and function, rather than merely naming organs or stating their basic functions (Lira & Gardner, 2020).

In the sperm pathway and reproductive disorder subtopics, differences in response quality became more evident. The experimental class was able to develop responses from blank answers into explanations that included content, prediction, and explanation, even accompanied by visualization in the sperm pathway subtopic (Ariely et al., 2024). In contrast, the control class still showed more fragmented mental models, especially when students were asked to explain mechanisms or cause-and-effect relationships. This finding indicates that Collaborative Constructive learning not only improved the completeness of students' responses, but also helped them organize knowledge into a more coherent line of reasoning (Prayitno et al., 2022). The use of visualization in science learning is also important because it can represent relationships among components and processes that are difficult to understand through verbal explanation alone (Bobek & Tversky, 2016).

Nevertheless, the examples of students' responses show that improvement in the experimental class was not evenly distributed across all aspects. Some responses were already strong in

content, prediction, and explanation, but were not yet accompanied by accurate visual representations, particularly in the reproductive organ and reproductive disorder subtopics (Andariana et al., 2020). This indicates that the Drawing aspect remained an obstacle in the formation of complete mental models. This finding is consistent with Nasution and Khairuna (2022), who indicated that reproductive system disorders are among the subtopics most prone to misconceptions. In Table 1.3, changes in the mental models of students in the experimental class were reflected not only in higher levels, but also in the improved quality of relationships among concepts in students' responses, especially in the Content, Prediction, and Explanation aspects (Batlolona & Diantoro, 2023).

Overall, the findings of this study indicate that the improvement of students' mental models in the experimental class can be explained through the principles of social constructivism (Mishra, 2023). Collaborative Constructive learning provides space for students to express ideas, compare understandings, and revise concepts through interactions with peers and the teacher. This process is important because mental models are not formed merely through the reception of information, but through active activities of organizing, connecting, and testing concepts. With the support of student worksheets, group discussion, and partially completed concept maps, students received scaffolding that helped them transform previously fragmented knowledge into a more directed structure of understanding (Mustafa et al., 2021).

However, the persistence of partial mental models among some students indicates that reconstruction of understanding had not occurred evenly across all C-P-E-D aspects. In several subtopics, students were able to master content and make predictions, but remained weak in constructing mechanistic explanations or visual representations. This condition may occur because the topic of the male reproductive system requires the integration of anatomical structure, physiological function, process sequencing, and cause-and-effect relationships. Students who do not yet have a strong conceptual foundation tend to construct fragmented understanding. This is in line with Belland (2017), who emphasized that scaffolding is necessary in STEM learning to help students deal with complex cognitive tasks.

The pedagogical implication of these findings is that biology learning, particularly in organ system topics, needs to be designed to develop systems thinking rather than merely emphasizing mastery of terminology. Teachers need to facilitate activities that help students connect structure, function, process, and disorders within a single sequence of biological reasoning. Strategies such as problem-based discussions, partially completed concept maps, exercises in explaining mechanisms, and gradual visual representation can be used to strengthen the Content, Prediction, Explanation, and Drawing aspects in a more balanced manner. Therefore, mental model assessment can serve as an alternative form of process assessment that helps teachers identify which parts of students' understanding remain partial and determine more appropriate forms of intervention.

Conclusion

This study demonstrates that the Collaborative Constructive approach has a positive effect on improving the quality of students' mental models in the topic of the male reproductive system. The experimental class showed greater improvement than the control class, as indicated by both the statistical test results and the shifts in mental model categories. Collaborative Constructive learning helped students develop more structured understanding, particularly in the aspects of Content, Prediction, and Explanation. However, some students still demonstrated partial mental models, especially in the Drawing aspect, indicating that their ability to visually represent anatomical structures and biological mechanisms still requires further strengthening.

The pedagogical implication for biology teachers is the need to design learning activities that do not merely emphasize mastery of terminology, but also train students to connect structure, function, process, and disorders within a coherent sequence of biological reasoning. Strategies such as guided group discussions, partially completed concept maps, exercises in explaining mechanisms, and gradual visual representation can be used to help students construct more complete mental models. Thus, teachers play a role not only as transmitters of content, but also as facilitators who provide scaffolding according to students' conceptual difficulties.

From an assessment perspective, the findings of this study emphasize that biology learning evaluation should be complemented by mental model-based assessment. The C-P-E-D framework can help teachers identify whether students' difficulties lie in content mastery, prediction ability, explanation construction, or visual representation. Future research is recommended to examine the sustainability of mental model changes through longitudinal studies and to develop more specific scaffolding strategies to strengthen the Drawing and Explanation aspects in other abstract and complex biological topics.

References

- Anastasiou, D., Wirngo, C. N., & Bagos, P. (2024). The Effectiveness of Concept Maps on Students' Achievement in Science: A Meta-Analysis. In *Educational Psychology Review* (Vol. 36, Issue 2). Springer US. <https://doi.org/10.1007/s10648-024-09877-y>
- Andariana, A., Zubaidah, S., Mahanal, S., & Suarsini, E. (2020). Identification of biology students' misconceptions in human anatomy and physiology course through three-tier diagnostic test. *Journal for the Education of Gifted Young Scientists*, 8(3), 1071–1085. <https://doi.org/10.17478/jegys.752438>
- Antinozzi, C., Di Luigi, L., Sireno, L., Caporossi, D., Dimauro, I., & Sgrò, P. (2025). Protective role of physical activity and antioxidant systems during spermatogenesis. *Biomolecules*, 15(4), 478. <https://doi.org/10.3390/biom15040478>
- Ariely, M., Nazaretsky, T., & Alexandron, G. (2024). Causal-mechanical explanations in biology: Applying automated assessment for personalized learning in the science classroom. *Journal of Research in Science Teaching*, 61(8), 1858–1889. <https://doi.org/10.1002/tea.21929>
- Baharuddin, R., & Umar, M. F. (2026). Assessing High School Students' Misconceptions of the Human Reproductive System Using a Three-Tier Diagnostic Test. *Pro-Life*, 13(2), 186-203. <https://doi.org/10.33541/pro-life.v13i2.8236>
- Batlolona, J. R., & Diantoro, M. (2023). MENTAL MODELS AND CREATIVE THINKING SKILLS IN STUDENTS' PHYSICS LEARNING. *Creativity Studies*, 16(2), 433–447. <https://doi.org/10.3846/cs.2023.14743>
- Belland, B. R. (2017). *Instructional Scaffolding in STEM Education*. Springer. <https://doi.org/10.1007/978-3-319-02565-0>
- Bobek, E., & Tversky, B. (2016). Creating visual explanations improves learning. *Cognitive Research: Principles and Implications*, 1(1), 1–14. <https://doi.org/10.1186/s41235-016-0031-6>
- Chiu, M. C., & Hwang, G. J. (2026). Enhancing student creative and critical thinking in generative AI-empowered creation: A mind-mapping approach. *Interactive Learning Environments*, 34(2), 869-890. <https://doi.org/10.1080/10494820.2025.2511244>

- Dong, Y., Liu, X., & Chen, Y. (2026). Exploring Pathways for Integrating Sex Education into High School Biology and Its Practical Significance. *Academic Journal of Education*, 1(1), 18-24. <https://doi.org/10.54097/b8y5mj20>
- Eun, B. (2017). The zone of proximal development as an overarching concept : A framework for synthesizing Vygotsky ' s theories The zone of proximal development as an overarching concept: A framework for synthesizing Vygotsky ' s theories. *Educational Philosophy and Theory*, 1857, 1–13. <https://doi.org/10.1080/00131857.2017.1421941>
- Gregorcic, T., & Torkar, G. (2022). Using the structure-behavior-function model in conjunction with augmented reality helps students understand the complexity of the circulatory system. *Advances in Physiology Education*, 46(3), 367–374. <https://doi.org/10.1152/advan.00015.2022>
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Hamdiyati, Y., Sudargo, F., Redjeki, S., & Fitriani, A. (2018). Perubahan model mental mahasiswa pada pokok bahasan virus melalui program perkuliahan mikrobiologi berbasis model mental. *EDUSAINS*, 10(1), 74–82. <https://doi.org/10.15408/es.v10i1.7777>
- Johnson, S., Walker, S. E., Goulding, C. J., LoTemplio, S., & Zarestky, J. (2026). “Nature from My Perspective”: Leveraging Participatory Mental Modeling to Understand the Well-Being Impacts of Outdoor Retreats for Black Men. *Ecopsychology*, 18(1), 56-72. <https://doi.org/10.1177/19429347251407579>
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge University Press.
- Kurnaz, M. A., & Eksi, C. (2015). An Analysis of High School Students ' Mental Models of Solid Friction in Physics. *Kuram ve Uygulamada Egitim Bilimleri*, 15(3), 787–795. <https://doi.org/10.12738/estp.2015.3.2526>
- Lata, P., & Hajam, Y. A. (2026). Introduction to male reproductive disorders. *Nanotherapeut Treat Reprod Disord*, 1-23.
- Lin, F., Ma, L., & Sheng, Z. (2025). Health disorders in menopausal women: microbiome alterations, associated problems, and possible treatments. *BioMedical Engineering OnLine*, 24(1), 84. <https://doi.org/10.1186/s12938-025-01415-3>
- Lin, Y. R., & Hung, C. Y. (2026). GAI-assisted self-regulation in science learning: the impact of interaction interfaces and content abstraction on flow and mental model construction. *Education and Information Technologies*, 31(10), 3369-3398. <https://doi.org/10.1007/s10639-026-13921-5>
- Lira, M., & Gardner, S. M. (2020). Leveraging Multiple Analytic Frameworks to Assess the Stability of Students ' Knowledge in Physiology. *CBE-Life Sciences Education*, 19(1), 1–19. <https://doi.org/10.1187/cbe.18-08-0160>
- Mishra, N. R. (2023). Constructivist Approach to Learning: An Analysis of Pedagogical Models of Social Constructivist Learning Theory. *Journal of Research and Development*, 6(01), 22–29. <https://doi.org/10.3126/jrdn.v6i01.55227>

- Momsen, J., Speth, E. B., Wyse, S., & Long, T. (2022). Using Systems and Systems Thinking to Unify Biology Education. *CBE-Life Sciences Education*, 21(2), 1–11. <https://doi.org/10.1187/cbe.21-05-0118>
- Mustafa, M., Ioannidis, A., Ferreira González, L., Dabrowski, T., & Großschedl, J. (2021). Fostering Learning with Incremental Scaffolds During Chemical Experimentation: A Study on Junior High School Students Working in Peer-Groups. *International Journal of Innovation in Science and Mathematics Education*, 29(2). <https://doi.org/10.30722/IJISME.29.02.002>
- Nasution, R. A., & Khairuna. (2022). Kombinasi CRI (Certainty of Response Index) dan Tes Pilihan Berganda dalam Menguji Tingkat Miskonsepsi Siswa Pada Materi Biologi Sistem Reproduksi Manusia. *Jurnal Pembelajaran Dan Biologi Nukleus*, 8(2), 373–385. <https://doi.org/10.36987/jpbn.v8i2.2858>
- Neway, A. (2026). *Enhancing Grade 11 Students' Conceptual Understanding of Genetics through Technology-Enhanced Guided Inquiry-Based Learning* (Doctoral dissertation).
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press.
- Prayitno, B. A., Sugiharto, B., & Titikusumawati, E. (2022). Effectiveness of Collaborative Constructivist Strategies to Minimize Gaps in Students' Understanding of Biological Concepts. *International Journal of Emerging Technologies in Learning (IJET)*, 17(11), 114–127. <https://doi.org/10.3991/ijet.v17i11.29891>
- Rachman, G., Wenno, I. H., Batlolona, J. R., Jamaludin, J., & Dulhasyim, A. B. P. (2026). An empirical and theoretical investigation using PBL: Students' mental models in understanding capillary rise. *Momentum: Physics Education Journal*, 10(1), 16-26. <https://doi.org/10.21067/mpej.v10i1.13143>
- Salsabilla, F., & Melta, D. (2024). Misconceptions Among High School Students Regarding The Biological Concepts of Human Reproduction 1. *Bioedunis Journal*, 3(2), 163–176. <https://doi.org/10.24952/bioedunis.v3i2.13876>
- Sanchez, T., Vereschak, O., & Deroy, O. (2026, March). Mental models in human-AI interaction: Systematic review of empirical methodologies and guidelines. In *Proceedings of the 31st International Conference on Intelligent User Interfaces* (pp. 663-682). <https://doi.org/10.1145/3742413.3789223>
- Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2017). Studying and Constructing Concept Maps : a Meta-Analysis. *Educational Psychology Review*, 30(2), 431–455. <https://doi.org/10.1007/s10648-017-9403-9>
- Seel, N. M. (2017). Model-based learning : a synthesis of theory and research. In *Educational Technology Research and Development*. Springer US. <https://doi.org/10.1007/s11423-016-9507-9>
- Shetty, P. K., Dindalkoppa, M., Rao, S. S., Pilli, G. S., & Sudhakar, S. (2025). Perception of case-based collaborative learning in pathology through panel discussion using concept mapping: a quasi-experimental crossover study in phase II MBBS students. *BMC Medical Education*, 26(1), 51. <https://doi.org/10.1186/s12909-025-08334-w>
- Sholihah, M. A., Zubaidah, S., Mahanal, S., & Listyorini, D. (2025). The effect of reading-concept mapping-reciprocal teaching on students' communication skills. *Journal of*

Education and Learning (EduLearn), 19(1), 158-168.
<https://doi.org/10.11591/edulearn.v19i1.21765>

- Stavros, S., Kathopoulis, N., Moustakli, E., Potiris, A., Anagnostaki, I., Topis, S., ... & Zikopoulos, A. (2025). Endocrine-disrupting chemicals and Male infertility: mechanisms, risks, and regulatory challenges. *Journal of Xenobiotics*, 15(5), 165. <https://doi.org/10.3390/jox15050165>
- Sunarto, M. D. (2025). Integrating AI in individualized mathematics learning: A teacher-centered approach. *Indonesian Journal of Science and Mathematics Education*, 8(3), 531-548.
- Vergara-Diaz, C., Navarrete, K., Parraguez, C., & Cofré, H. (2026, March). Mental models of the earth's internal structure among primary, middle and secondary school students in Chile. In *Frontiers in Education* (Vol. 11, p. 1737386). Frontiers Media SA. <https://doi.org/10.3389/feduc.2026.1737386>
- Wahyuni, N., Suwono, H., & Lestari, U. (2019). Learning Difficulties of High School Students in Understanding. *Jurnal Pendidikan Sains*, 7(4), 116–121.
- Walton, K. L. W. (2023). Use of a short , in-class drawing activity to assess student understanding of core concepts of the cell membrane in an undergraduate physiology course. *Advances in Physiology Education*, May 2023, 508–513. <https://doi.org/10.1152/advan.00218.2022>
- Zhang, X., Zhang, L., & Xiang, W. (2025). The impact of mitochondrial dysfunction on ovarian aging. *Journal of translational medicine*, 23(1), 211. <https://doi.org/10.1186/s12967-025-06223-w>