education sciences

MDPI

Article

Enhancing Teaching Strategies through Cognitive Load Theory: Process vs. Product Worked Examples

Gerry Sozio [1,*], Shirley Agostinho [2], Sharon Tindall-Ford [2] and Fred Paas [2,3]

[1] Education and Research Directorate, Sydney Catholic Schools, Sydney, NSW 2000, Australia
[2] School of Education, University of Wollongong, Wollongong, NSW 2522, Australia; shirleya@uow.edu.au (S.A.); sharontf@uow.edu.au (S.T.-F.); paas@essb.eur.nl (F.P.)
[3] Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, 3062 PA Rotterdam, The Netherlands
* Correspondence: gerardo.sozio@syd.catholic.edu.au; Tel.: +61-402-097-461

Abstract: Based on Cognitive Load Theory (CLT), learning by studying worked examples, that is, step-by-step solutions, has been shown to reduce cognitive load and enhance learning outcomes more than learning by solving conventional problems. Two types of worked examples that have been predominantly used in well-structured learning domains include product-oriented worked examples (ProductWEs) which provide strategic information and process-oriented worked examples (ProcessWEs) which provide strategic information and principled knowledge. However, less research has been conducted on worked examples' effectiveness in ill-structured learning domains. In a study with 85 university students enrolled in teaching programs, we investigated whether ProcessWEs or ProductWEs better support identifying and applying knowledge regarding the quality teaching component of substantive communication. Participants completed tasks under three instructional conditions: ProcessWE, ProductWE, and conventional problem-solving. Results showed that ProcessWEs outperformed ProductWEs, ProductWEs outperformed conventional problem-solving, and conventional problem-solving had higher perceived task difficulty than the other two conditions. This study theoretically contributes to CLT research as it found that the use of ProcessWEs and ProductWEs is effective in an ill-structured learning domain and makes a practical contribution by showing what a ProductWE or ProcessWE in an ill-structured learning domain 'looks like'.

Keywords: Cognitive Load Theory (CLT); worked examples; process-oriented worked examples (ProcessWEs); product-oriented worked examples (ProductWEs); mental effort; task difficulty; New South Wales Quality Teaching Model (NSW QTM); substantive communication

check for updates

Citation: Sozio, G.; Agostinho, S.; Tindall-Ford, S.; Paas, F. Enhancing Teaching Strategies through Cognitive Load Theory: Process vs. Product Worked Examples. *Educ. Sci.* **2024**, *14*, 813. https://doi.org/10.3390/educsci14080813

Academic Editor: Don Passey

Received: 20 May 2024
Revised: 22 July 2024
Accepted: 23 July 2024
Published: 25 July 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

An important focus of Cognitive Load Theory (CLT) has been on the effectiveness of worked examples in well-structured learning domains with clearly defined problems and solution methods [1]. Worked examples offer step-by-step solutions, aiding learners in acquiring problem-solving skills [1] and modeling to learners how to perform a particular skill or task [2]. Worked examples reduce cognitive load compared to conventional problem-solving [1] because the steps-to-solution as provided in ProductWEs and steps-to-solution with an explanatory rationale as provided in ProcessWEs help to facilitate efficient use of working memory [3,4]. Most CLT research has been conducted in well-structured learning domains such as mathematics, where there are procedural steps to achieve one correct solution [1,5]. The effectiveness of worked examples in ill-structured learning domains, prevalent in humanities and the arts, remains underexplored [1,5]. The purpose of this research was to investigate how worked examples would transfer to an ill-structured learning domain, that is, a topic or discipline that is more interpretative in solution [6]. The domain of the worked example in this research is quality teaching.

This research was underpinned by CLT, an instructional theory developed in the 1980s [3]. CLT frames learning as an information processing system involving working memory (WM) and long-term memory (LTM). CLT states that for learning to be optimized, knowledge needs to be processed in WM, then stored in LTM, then accessed and transferred back into WM to make sense of and process new information. WM has a very limited capacity [7–9] and is used to process information, which is then transferred into LTM [1,10]. For this to occur, limited WM resources need to be optimized. Instructional designs based on CLT principles enable improved processing abilities for learners through optimizing WM capacity [4,11,12].

CLT identifies two types of cognitive load imposed during information processing in WM: intrinsic and extraneous cognitive load [1]. Intrinsic cognitive load (ICL) is determined by the complexity of the material, which is based on the number of interactive information elements requiring simultaneous processing. Element interactivity can be determined by simultaneously considering the structure of the information being processed and the knowledge held in long-term memory of the person processing the information, i.e., prior knowledge [13,14]. Extraneous cognitive load (ECL) relates to how instructional materials are presented, causing cognitive activities unrelated to learning and increasing element interactivity [1,4]. This in turn increases element interactivity, which places an additional burden on limited WM resources [4,11,14,15]. Reducing ECL becomes crucial when ICL is high, as the two loads are additive [1].

CLT has played a pivotal role in developing various instructional designs, known as CLT effects, aimed at enhancing learning [4]. This study primarily drew on the worked example effect, while also considering other effects such as the split attention effect and the redundancy effect, specifically in relation to the design of the research instructional material. Split attention occurs when learners are required to process multiple sources of information simultaneously to make meaning [10], and the redundancy effect occurs when unnecessary visual or auditory elements are presented in learning content and cause additional ECL on WM, thus impacting learning [1].

The worked example effect emerged in mathematics, a well-structured learning domain with clearly defined problems and solutions [1,16]. Worked examples, which present problem states followed by detailed, step-by-step solutions, help reduce cognitive load by leading learners through a forward problem-solving strategy like that of experts [17,18]. This approach contrasts with the less efficient means–ends strategy often used by novices, which can overload working memory. The means–ends approach involves problem-solvers working backwards from the goal, establishing intermediate objectives until they reach a point where all remaining steps are clear and straightforward, leading directly to the desired outcome. The worked example models an expert-like approach to problem-solving, and reduces cognitive load through promoting a forward working strategy [18].

Process-oriented worked examples (ProcessWEs) and product-oriented worked examples (ProductWEs) were investigated in this research. ProcessWEs offer not only strategic information, that is, the solution steps demonstrating how to solve the problem, but also the underlying rationale, integrating principled knowledge that benefits novice learners. In contrast, ProductWEs present the solution steps demonstrating how to solve the problem (strategic information) without explaining the rationale behind them [19]. Research suggests that ProcessWEs enhance cognitive schema construction and automation, leading to better performance compared to ProductWEs [13,19–22]. Moreover, inclusion of principled knowledge in ProcessWEs improves understanding by encouraging learners to invest effort in understanding "why" and "how" [23].

2. The New South Wales Quality Teaching Model

This research examined the effectiveness of ProcessWEs and ProductWEs in aiding pre-service teachers' comprehension and implementation of the New South Wales Quality Teaching Model (NSW QTM) [24,25], particularly focusing on substantive communication. The NSW QTM offers a comprehensive framework for assessing teaching quality. Its appli-

cation is considered an ill-structured learning domain due to the absence of clearly defined problem states, solution pathways, and predetermined correct answers [1,6]. Despite this complexity, the NSW QTM has been effectively employed to support professional learning in schools through Quality Teaching Rounds [26]. Quality Teaching Rounds involve teams of teachers observing each other's lessons and collaboratively evaluating the teaching using the NSW QTM. This process fosters in-depth discussions and reflections on teaching practices [26]. The implementation of this model has been shown to enhance teacher professional growth, elevate teaching standards, inspire teacher motivation, and foster a positive school environment [26].

The NSW QTM consists of three dimensions: intellectual quality, quality learning environment, and significance. Each of these dimensions encompasses six specific elements. The details of these dimensions and their respective elements are presented in Table 1.

Table 1. Dimensions and elements of the Quality Teaching Model [26].

Quality Learning Environment	Intellectual Quality	Significance
Explicit Quality Criteria	Deep Knowledge	Background Knowledge
Engagement	Deep Understanding	Cultural Knowledge
High Expectations	Problematic Knowledge	Knowledge Integration
Social Support	Higher Order Thinking	Inclusivity
Students' Self-regulation	Metalanguages	Connectedness
Student Direction	Substantive communication	Narrative

Substantive communication, as defined by the New South Wales Department of Education and Training (2003), involves students engaging in sustained and meaningful discussions about the concepts and ideas they encounter, emphasizing the quality of communication necessary to foster understanding. These discussions, occurring between both teachers and students, as well as among students themselves, can take various forms such as oral, written, or artistic expressions [24,25].

The NSW QTM outlines the key characteristics of lessons characterized by high levels of substantive communication:

- There is sustained interaction: The continuation of a thought, idea, or concept beyond the initiate–respond–evaluate (IRE) method of questioning, where there is a sustained coherent flow of communication where the flow of ideas may include extended statements, questions, or statements from one person to another [24,25].
- The communication is focused on the substance of the lesson: The communication moves beyond recounting facts, definitions, experiences, and procedures and needs to promote critical reasoning [24,25].
- The interaction is reciprocal, that is, two-way: There is a flow of ideas that is at least two-way in direction [24,25].

Substantive communication was chosen as a focal point because it is a crucial component of the NSW QTM, stimulating thinking, engagement, and discussions in the classroom to promote learning [24,25]. Furthermore, the authors consider that understanding and implementing substantive communication within a classroom setting is complex for both novices and experts, ensuring the content is high in element interactivity. This complexity arises from the need to grasp the three characteristics of substantive communication and retain them in working memory to assess its level accurately. The challenge is even more significant for novices due to their limited prior knowledge and lack of teaching experience [13].

3. Materials and Methods

This study aimed to determine the most effective type of worked example for supporting novice learners' understanding within an ill-structured learning domain. Although the use of ProcessWEs and ProductWEs has been researched and validated in well-structured

learning domains such as mathematics, science, and economics [1,27–29], less attention has been given to their effectiveness in ill-structured learning domains [1,5]. This study sought to address this research gap by evaluating the effectiveness of both types of worked examples in such a learning domain. The experiment employed a quantitative research experimental design to investigate the effectiveness of ProcessWEs and ProductWEs for pre-service teachers (novices) learning about substantive communication characteristics within an ill-structured learning domain. The ProcessWEs and ProductWEs used in this study included online video recordings of lessons being taught, annotations on the video recordings indicating whether substantive communication was present or not during the section of the lesson that the participants had watched, and hard copy documents. The hard copy documents included referencing tables indicating the level of substantive communication that was demonstrated during the video recording of the lesson which had just been watched. Further information on the ProcessWEs and ProductWEs used in this study is presented in Section 3.5.2.

As stated above, the instructional content used in this study was based on the NSW QTM, specifically, the element substantive communication as it is an important component of the NSW QTM aiming to foster classroom engagement, discussions, and learning [24,25]. Substantive communication was chosen as understanding how it can be enacted within a classroom is complex for novice teachers, thus ensuring the content was high in element interactivity. Learning to understand and apply the NSW QTM demonstrates the characteristics of an ill-structured learning domain. It does not involve specifying the various problem states and problem-solving operators. Moreover, it does not involve providing a specific, pre-determined correct answer [1,6].

3.1. Research Questions and Related Hypotheses

The overall aim of this research was to investigate which instructional condition supported participants best in identifying and applying knowledge of the New South Wales Quality Teaching Model element of substantive communication through considering participants' cognitive load and perceived task difficulty.

The theoretical justification for Hypotheses 1 and 2 is based on CLT, which emphasizes optimizing WM capacity for effective learning. ProcessWEs provide both solution steps and underlying rationale [20], reducing intrinsic and extraneous cognitive load, thus facilitating better knowledge transfer to long-term memory and enhancing test performance [13,19–22]. ProductWEs, which offer only the solution steps [19], also reduce extraneous cognitive load compared to unguided problem-solving in the Control condition, leading to better test performance by efficiently managing WM resources. Consequently, participants in the ProcessWE condition are expected to outperform those in both the ProductWE and Control conditions on test scores. Additionally, participants in the ProductWE condition are expected to achieve higher test scores than those in the Control condition [13,19–22]. The research questions and associated hypotheses for the experiment were the following:

Research Question 1: When learning about the characteristics of substantive communication, do participants presented with the ProcessWE condition achieve higher test performance scores during the test phase than participants presented with the ProductWE and Control conditions, and do participants presented with the ProductWE condition achieve higher test performance scores during the test phase than participants presented with the Control condition?

Hypothesis 1. *The ProcessWE condition will outperform the ProductWE and Control conditions on test phase scores.*

Hypothesis 2. *The ProductWE condition will outperform the Control condition on test phase scores.*

Research Question 2: When learning about the characteristics of substantive communication, do participants presented with the ProcessWE condition or Control condition report higher perceived cognitive load and higher perceived task difficulty during the learning phase than participants presented with the ProductWE condition during the learning phase?

Hypothesis 3. *The ProcessWE condition or Control condition will report higher perceived cognitive load than the ProductWE condition during the learning phase.*

Hypothesis 4. *The ProcessWE condition or Control condition will report higher perceived task difficulty than the ProductWE condition during the learning phase.*

Research Question 3: When learning about the characteristics of substantive communication, do participants presented with the ProcessWE condition report lower perceived cognitive load and lower perceived task difficulty during the test phase than participants presented with the ProductWE condition?

Hypothesis 5. *The ProcessWE condition will report lower perceived cognitive load than the ProductWE condition during the test phase.*

Hypothesis 6. *The ProcessWE condition will report lower perceived task difficulty than the ProductWE condition during the test phase.*

Research Question 4: When learning about the characteristics of substantive communication, do participants presented with the ProcessWE condition or ProductWE condition report lower perceived cognitive load and lower perceived task difficulty during the test phase than participants presented with the Control condition?

Hypothesis 7. *The ProcessWE condition or ProductWE condition will report lower perceived cognitive load than the Control condition during the test phase.*

Hypothesis 8. *The ProcessWE condition or ProductWE condition will report lower perceived task difficulty than the Control condition during the test phase.*

3.2. Research Design

This study used a between-subjects design to examine test performance scores, perceived mental effort, and perceived task difficulty ratings of participants when learning about and completing NSW QTM substantive communication tasks. A one-way Analysis of Variance (ANOVA) between instructional conditions was conducted for each of the dependent variables. The independent variable was the instructional condition, that is, the ProcessWE condition, ProductWE condition, or Control condition. The dependent variables were as follows:

1. Performance on test items.
2. Reported perceived mental effort measured by the Cognitive Load Subjective Rating Scale [30].
3. Reported perceived task difficulty measured by the Task Difficulty Subjective Rating Scale [31,32].

The Cognitive Load Subjective Rating Scale, developed by Paas in 1992 [30], is a nine-point scale assessing participants' perceived mental effort, validated for measuring cognitive load. Similarly, the Task Difficulty Subjective Rating Scale, also a nine-point scale, evaluates participants' perceived task difficulty [31]. Figures 1 and 2 below present both scales:

Rate the amount of mental effort you invested to complete the task.

1	2	3	4	5	6	7	8	9
Extremely low mental effort				Neither low nor high mental effort				Extremely high mental effort

Figure 1. Representation of the Cognitive Load Subjective Rating Scale [30].

Rate how difficult the task was for you.

1	2	3	4	5	6	7	8	9
Extremely easy				Neither easy nor difficult				Extremely difficult

Figure 2. Representation of the Task Difficulty Rating Scale [31].

Although perceived task difficulty and perceived mental effort may correlate, they represent distinct constructs [31]. Mental effort refers to the total controlled cognitive processing exerted by a subject, often described as "how hard I tried" [33] (p. 1), while perceived task difficulty primarily concerns the task itself, described as "how difficult the task is" [33] (p. 1). Throughout the experiment, participants rated their perceived mental effort and task difficulty after each task in both the learning and test phases.

3.3. Participants

Participants included pre-service teachers enrolled in either their first year of a two-year post-graduate Master of Teaching program (this program offers both primary and secondary school focus) or their final year of a four-year Bachelor of Primary Education program at an Australian regional university. There were 55 Master of Teaching participants (18 male, 37 female) aged 21 to 51 years, and 30 Fourth Year Bachelor of Primary Education participants (13 male, 17 female) aged 22 to 38 years. As pre-service teachers have limited understanding and application of the NSW QTM and substantive communication due to their training status, both Master of Teaching and Bachelor of Primary Education participants were considered novices. Table 2 summarizes participant numbers, age ranges, and allocations across the instructional conditions of ProcessWE and ProductWE and the Control condition of conventional problem-solving. The lower number of participants in the Control condition was because it was scheduled as the final session in the experiment.

Table 2. Range of ages and number of male and female participants for the three instructional conditions.

Range of Ages and Number of Male and Female Participants					
Instructional Condition	**Range of Ages (Years)**	**Males**	**Females**	**Total**	**Total of Master of Teaching/Bachelor of Primary Education Participants**
ProcessWE	21–26	11	29	40	27/13
ProductWE	22–47	15	17	32	18/14
Control Condition	21–51	5	8	13	10/3

3.4. Procedure

During the experiment, participants received two worked examples tailored to their instructional condition. These examples comprised video recordings of Personal Development, Health and Physical Education (PDHPE) and Japanese language lessons, along with

referencing tables indicating the level of substantive communication and other instructional condition-specific information, including annotations on the videos and justifications of the levels of substantive communication on the referencing table. Participants were trained to identify and rate levels of substantive communication during lessons. The experiment comprised three phases: introductory (15 min), learning (25 min), and test (20 min), totaling 60 minutes.

During the introductory phase, the researcher ensured participants understood the experiment's structure, purpose, and tasks. Procedures were explained consistently, tailored to each instructional condition. Participants were informed they would complete test items and rate their perceived mental effort and perceived task difficulty afterward. The researcher provided an example of rating mental effort and task difficulty. Participants learned about low levels of substantive communication and were introduced to Coding, allowing them to rate substantive communication effectiveness using a five-point scale. A rating of one indicates low substantive communication, while a rating of five indicates high substantive communication.

In the learning phase, participants were instructed to begin with page one of their booklet. They were informed they would watch two video recordings of lessons, serving as worked examples to understand the characteristics and levels of substantive communication and apply this knowledge to teaching practices. The video recordings were shown on a large screen, with ten-second pauses at key points in each lesson for all instructional conditions. Annotations in the ProcessWE and ProductWE conditions indicated the level of substantive communication prior to each pause. While the Control condition also had pauses, they had no explanatory annotations. These pauses ensured equal duration for all recordings and allowed participants in each condition to reflect on the level of substantive communication uniformly.

Following the learning phase, participants proceeded to the test phase, which included three test items and an evaluation in the participant booklet. After completing each test item, participants had twenty seconds to rate their perceived mental effort and task difficulty. Following the completion of the three test items, participants evaluated their experience in the experiment; this involved rating the following statements, "I enjoyed learning in this way" and "I found this type of instruction engaging" on a five-point Likert scale.

3.5. Instructional Materials

The experimental materials were developed by the researcher in consultation with PhD supervisors, international expert CLT researchers, and three Education Officers from a regional Catholic Education Diocese. Further advice from CLT experts was obtained during presentations of the details of the experiment at the International Cognitive Load Theory Conferences in the United States of America in 2015, Germany in 2016, and Australia in 2017.

Participants were presented with one of three instructional conditions:

1. ProcessWE condition.
2. ProductWE condition.
3. Control condition.

The ProcessWEs and ProductWEs used in the experiment included online video recordings of lessons being taught, annotations on the video recordings, and hard copy documents. The participants in the Control condition were provided with practice problems to solve without any support or guidance during the practice activity. The experiment consisted of three phases:

1. Introductory phase (Phase 1).
2. Learning phase (Phase 2).
3. Test phase (Phase 3).

3.5.1. Phase 1—PowerPoint Presentation: Introductory Phase (15 min)

During the introductory phase (Phase 1), participants in each instructional condition were presented with a 16-slide Microsoft (MS) PowerPoint presentation introducing the NSW QTM and focusing on substantive communication. The slides covered an overview of the NSW QTM, a definition of substantive communication, characteristics associated with high and low levels of substantive communication, and the Coding Scale used to assess substantive communication effectiveness. Additional slides included screen captures of annotations in the video recordings participants would watch during the learning phase.

3.5.2. Phase 2—Worked Examples: Learning Phase (25 min)

The worked examples consisted of video recordings of the lessons along with the referencing tables included in the accompanying hard copy participant booklet. Video recordings of the lessons were developed to reflect the three instructional conditions of this study: ProcessWE condition, ProductWE condition, and Control condition.

Video Recordings of the Lessons

This study utilized two video recordings produced by the New South Wales Department of Education and Training [34,35] to aid teachers in comprehending the NSW QTM's 18 elements, with a focus on substantive communication. The first recording captured a PDHPE lesson, lasting seven minutes and forty-six seconds, addressing safety issues among middle school female students. The second recording featured a Japanese language lesson lasting four minutes and fifteen seconds, teaching counting in Japanese to a mixed group of middle school students. A specific process was followed by the researcher to adapt these recordings for the three instructional conditions (ProcessWE, ProductWE, and Control):

- Video recordings were analyzed for levels of substantive communication and were edited to include annotations relevant to the instructional condition.
- Scripts were written to be delivered to participants for each instructional condition.
- Ten-second pauses were embedded at key points during the video recording of the lessons for each instructional condition (this was so the video recordings for each instructional condition were the same length).

The author used information from accompanying resources provided by NSW DET [35]. Validation was sought from three Education Officers employed at a Catholic Education Office, responsible for supporting a system of schools, to confirm the identified levels of substantive communication featured in the annotations within the video recordings of the lesson.

Throughout the video recordings, ten-second pauses occurred at pivotal points, during which annotations were displayed on the screen for the two worked example instructional conditions (ProcessWE and ProductWE), denoting the level of substantive communication observed in the preceding section of the lesson. Additionally, the ProcessWE condition annotations included a rationale explaining the level of substantive communication. Conversely, no annotations were presented for the Control condition.

During the PDHPE lesson recording, there were sixteen ten-second pauses, and seven ten-second pauses during the Japanese language lesson recording. In the ProcessWE condition video recordings, annotations on the screen indicated the presence or absence of substantive communication, accompanied by concise descriptions explaining its level. Annotations for the ProcessWE consisted of up to three lines, including commentary and responses of teachers (T) and students (S). Each annotation remained visible for ten seconds to allow participants to assess substantive communication. Figure 3 presents an example of an annotation for the ProcessWE condition.

The ProductWE condition video recordings only included annotations on the screen at key points indicating whether substantive communication was evident or not during the section of the lesson that the participants had just watched, i.e., the annotations were 'SC evident' or 'SC not evident', no explanation of why substantive communication was

evident or not was provided, unlike the ProcessWE condition. The timing for the pauses was also ten seconds.

The Control condition video recordings did not include any annotations on the screen at key points indicating whether substantive communication was evident or not during the section of the lesson that the participants had just watched. The participants were notified that there would be pauses at key points during the video recordings and that during these pauses they were to reflect on the level of substantive communication within the lesson prior to the pause. The timing for the pauses was also ten seconds.



Figure 3. Example of an annotation provided during the video recording of a lesson for the ProcessWE condition [34,35].

Participation Booklet

A 12-page booklet was designed and included text material to accompany the video recordings of the lessons that provided steps to solutions through the referencing tables presented with the video recordings during the learning phase. Additionally, the booklet comprised test items, mental effort measures, task difficulty measures, and an evaluation/reflection section for participants to complete.

The participant booklet was designed to scaffold the lesson video recordings and facilitated the experimental process in Phases 2 and 3. It contained two worked examples (Tasks 2 and 3) presenting solutions, including Coding Scores reflecting the level of substantive communication observed in the videos for both ProcessWE and ProductWE conditions. The booklet also provided Coding Scores for the Control condition's video recordings, presented in a referencing table, with content tailored to each instructional condition.

The 12-page booklet, printed on single-sided A4 paper, displayed a cover page titled "The NSW Quality Teaching Model", with a section prompting participants to write their participation number, gender, and age. Participant numbers were assigned to the booklets before the experiment and distributed sequentially, starting with those seated at the front. The numbering system indicated the instructional condition: PP for the ProcessWE condition (e.g., PP12), PT for the ProductWE condition, and CT for the Control condition.

On the first page of the booklet, participants were instructed to respond to two statements aimed at gauging their prior knowledge of the NSW QTM and substantive communication. These responses helped confirm participants as novices. The statements were the following:

- Write down everything you know about the NSW Quality Teaching Model.
- Write down everything you know about substantive communication.

The researcher assessed these responses, revealing a mean score of 0.4/4 with a standard deviation of 0.8, confirming participants as novices regarding the NSW QTM and substantive communication.

The booklet included six tasks for participants, with the final task involving an evaluation. Tasks 1 and 2 focused on engaging with the worked examples during the learning phase. The two worked examples focused on a PDHPE and a language lesson and included a referencing table which provided a Coding Score for the level of substantive communication evident in that section of the lesson. Tasks 3, 4, and 5 comprised test items for the test phase. Each test item was on a separate page, followed by tables for mental effort rating and task difficulty rating. Task 6 involved participants evaluating their experiment experience.

3.5.3. Phase 3—Test Items: Test Phase (20 min)

The test phase followed immediately after the learning phase. The test phase required participants to complete three test items and an evaluation. The researcher designed test items comprising recall tasks, near transfer tasks, and far transfer tasks. Recall tasks involve recalling information, near transfer tasks require applying learnings to similar tasks, and far transfer tasks involve applying learnings to new situations. The paragraph below provides an overview of tasks presented in the participant booklet during the test phase, which comprised three tasks.

The first task included two sections, a recall task and a near transfer task for a total of three marks each. The task required participants to recall and identify the characteristics of substantive communication in a vignette of a conversation between a teacher and students. The second task included three sections, two near transfer tasks and a far transfer task for a total of one mark each. The near transfer task required participants to identify and justify characteristics of reciprocal communication, a characteristic of substantive communication, after watching a video of an interaction between a teacher and students. The third task included a far transfer task for a total of four marks. The task required participants to provide strategies on how they would enhance two identified characteristics of substantive communication after watching a video of an interaction between a teacher and students.

3.5.4. Mental Effort and Task Difficulty Ratings

After each task in the learning phase and subsequent test phase, a nine-point mental effort rating and task difficulty rating [30] were presented for participants to complete. These scales are presented in Figures 1 and 2.

4. Results

The dependent variables under analysis were individual test performance scores, total test performance scores, individual task perceived mental effort ratings, learning phase perceived mental effort ratings, test phase perceived mental effort ratings, individual perceived task difficulty ratings, learning phase perceived task difficulty ratings, test phase perceived task difficulty ratings, and evaluation ratings. Results for the dependent variables are presented in Sections 4.1–4.3.

4.1. Test Performance Scores

A one-way ANOVA was conducted on test performance scores to examine differences between the three instructional conditions. The homogeneity of variances test showed non-significance ($p > 0.05$), indicating homogeneous variances within conditions. Test performance scores, including mean and standard deviation for each condition, are presented in Table 3, with an asterisk denoting a significant main effect. Cohen's d was calculated for effect size, where values of 0.20, 0.50, and 0.80 represented small, medium, and large effects, respectively [36].

Results from the one-way ANOVA for test performance indicated a significant main effect for the test performance scores, $F(2, 82) = 4.60$, $p = 0.013$. Post hoc comparisons using the Tukey HSD test showed that the test performance scores for the ProcessWE

condition were significantly higher than the test performance scores for the ProductWE condition, $p = 0.02$, with a medium effect size obtained, $d = 0.64$. The Tukey HSD test is a post hoc test used to determine whether a set of conditions significantly differs from one or more others [37]. Comparing the ProcessWE and Control condition test performance scores did not yield statistically significant results, despite a higher mean score for the ProcessWE condition.

Table 3. Means and standard deviations for the test performance scores during the test phase.

	Test Performance Scores			
Instructional Condition	**Task 1 Recall Task (6)**	**Task 2 Near/Far Transfer Task (3)**	**Task 3 Far Transfer Task (4)**	**Total (13)**
ProcessWE (N = 40)	3.14 (1.28) *	2.28 (1.01)	1.83 (0.87)	7.24 (2.41) *
ProductWE (N = 32)	2.16 (1.55)	2.11 (1.05)	1.45 (1.04)	5.72 (2.34)
Control Condition (N = 13)	2.50 (1.49)	1.69 (1.18)	1.46 (0.97)	5.65 (2.02)

* Denotes significant main effect.

The one-way ANOVA revealed a significant main effect for Task 1 test performance scores, $F(2, 82) = 4.347$, $p = 0.016$. Post hoc Tukey HSD comparisons indicated significantly higher Task 1 scores for the ProcessWE condition compared to the ProductWE condition, $p = 0.013$. However, no significant main effects were found for Task 2, $F(2, 82) = 1.512$, $p = 0.227$, and Task 3, $F(2, 82) = 1.587$, $p = 0.211$.

Task 1 results and overall test scores confirmed Hypothesis 1, showing higher performance for participants in the ProcessWE condition compared to the ProductWE condition when learning about substantive communication characteristics. However, Hypothesis 2, indicating better performance for ProductWE condition participants over Control condition participants, did not yield statistically significant results, despite a higher mean score for the ProductWE condition.

4.2. Mental Effort Ratings

A one-way ANOVA analyzed perceived mental effort ratings to detect differences among the three instructional conditions. Participants assessed their perceived mental effort after viewing both worked examples in the learning phase and after each of the three test items in the test phase. The homogeneity of variances test showed non-significance ($p > 0.05$) for mental effort ratings during the learning phase, indicating homogeneous variances within instructional conditions. The result from the homogeneity of variances test was not significant ($p > 0.05$) for the mental effort ratings during the test phase; therefore, the variances within conditions were considered to be homogeneous.

Mental effort ratings, including mean and standard deviation for each condition during the learning and test phases, are presented alongside the test performance scores in Table 4, with an asterisk marking significant main effects.

The one-way ANOVA for mental effort ratings during the learning phase found no significant main effect for the total learning phase, $F(2, 82) = 2.346$, $p = 0.102$. Although non-significant, the Control condition showed a higher mean mental effort rating compared to the Product condition, with a medium effect size ($d = 0.67$). Similarly, the ProcessWE condition exhibited a higher mean mental effort rating than the ProductWE condition, with a medium effect size ($d = 0.26$).

Table 4. Means and standard deviations for the test performance scores and mental effort ratings during the learning and test phases.

	Test Performance Scores				Mental Effort Ratings Learning Phase			Mental Effort Ratings Test Phase			
Instructional Condition	**Task 1 Recall (6)**	**Task 2 Near/Far (3)**	**Task 3 Far (4)**	**Total (13)**	**Worked Example 1 (9)**	**Worked Example 2 (9)**	**Total (9)**	**Task 1 (9)**	**Task 2 (9)**	**Task 3 (9)**	**Task 4 (9)**
ProcessWE (N = 40)	3.14 (1.28) *	2.28 (1.01)	1.83 (0.87)	7.24 (2.41) *	3.08 (1.61)	2.62 (1.31)	2.89 (1.36)	5.23 (1.44)	4.90 (1.34)	5.95 (1.20)	5.36 (1.08)
Product WE (N = 32)	2.16 (1.55)	2.11 (1.05)	1.45 (1.04)	5.72 (2.34)	2.91 (1.87)	2.13 (1.76)	2.52 (1.51)	4.91 (1.67)	4.66 (1.72)	5.34 (1.70)	4.97 (1.29)
Control Condition (N = 13)	2.50 (1.49)	1.69 (1.18)	1.46 (0.97)	5.65 (2.02)	3.23 (1.64)	3.85 (1.95) *	3.54 (1.53)	5.46 (1.81)	5.38 (1.45)	6.38 (1.26)	5.74 (1.15)

* Denotes significant main effect.

The one-way ANOVA for mental effort ratings during the learning phase showed no significant main effect for Worked Example 1, $F(2, 82) = 0.185$, $p = 0.831$. However, there was a significant main effect for Worked Example 2, $F(2, 82) = 5.373$, $p = 0.01$. Post hoc comparisons revealed that the mental effort rating for Worked Example 2 in both the ProcessWE and ProductWE conditions was significantly lower than in the Control condition ($p = 0.004$ for the ProductWE condition and $p = 0.048$ for the ProcessWE condition). These results align with Hypothesis 3, indicating higher perceived mental effort for participants in the Control condition compared to the ProductWE condition when learning about substantive communication.

The one-way ANOVA for mental effort ratings during the test phase showed no significant main effect for the total test phase, $F(2, 82) = 2.214$, $p = 0.116$. However, the mean mental effort rating for the Control condition was higher than for both the ProcessWE and ProductWE conditions, with medium effect sizes ($d = 0.34$ for the ProcessWE condition and $d = 0.63$ for the ProductWE condition).

The one-way ANOVA for mental effort ratings during the test phase showed no significant main effect for Task 1, $F(2, 82) = 0.670$, $p = 0.514$, for Task 2, $F(2, 82) = 1.087$, $p = 0.342$, and for Task 3, $F(2, 82) = 2.995$, $p = 0.056$. Task 3, a far transfer task and the most complex of the tasks, elicited the highest mental effort ratings among participants, indicating their ability to accurately assess cognitive load.

Hypothesis 5, suggesting lower perceived mental effort for participants in the ProcessWE condition compared to the ProductWE condition during the test phase, lacked statistical significance in the experiment, thus not supporting the hypothesis. Similarly, Hypothesis 7, proposing lower perceived mental effort for participants in the ProcessWE or ProductWE conditions compared to the Control condition, did not achieve statistical significance. Despite lower mean mental effort ratings for the ProcessWE and ProductWE conditions, no significant difference was observed.

4.3. Task Difficulty Ratings

A one-way ANOVA was conducted on perceived task difficulty ratings to assess differences between the three instructional conditions. Participants rated their perceived task difficulty after each worked example presentation during the learning phase and after each task during the test phase. The homogeneity of variances test yielded significant results for the task difficulty ratings during the learning phase and for Task 2 during the test phase, indicating non-homogeneous variances within conditions. However, variances were considered homogeneous for Task 1, Task 3, and the total test phase. The result from the homogeneity of variances test was not significant ($p > 0.05$) for the task difficulty ratings during the test phase for all ratings except for Task 2; therefore, the variances within instructional conditions were considered to be homogeneous. Table 5 presents test performance scores alongside the perceived task difficulty ratings, including mean and

standard deviation for each instructional condition during the learning and test phases, with an asterisk denoting a significant main effect.

Table 5. Means and standard deviations for the test performance scores and task difficulty ratings during the learning and test phases.

	Test Performance Scores				Task Difficulty Ratings Learning Phase			Task Difficulty Ratings Test Phase			
Instructional Condition	**Task 1 Recall (6)**	**Task 2 Near/Far (3)**	**Task 3 Far (4)**	**Total (13)**	**Worked Example 1 (9)**	**Worked Example 2 (9)**	**Total (9)**	**Task 1 (9)**	**Task 2 (9)**	**Task 3 (9)**	**Task 4 (9)**
ProcessWE (N = 40)	3.14 (1.28) *	2.28 (1.01)	1.83 (0.87)	7.24 (2.41) *	1.98 (1.10)	1.87 (1.08)	1.91 (1.02)	4.20 (1.74)	3.78 (1.54)	5.40 (1.58)	4.46 (1.32)
Product WE (N = 32)	2.16 (1.55)	2.11 (1.05)	1.45 (1.04)	5.72 (2.34)	2.99 (1.48)	1.63 (1.04)	1.81 (0.97)	3.97 (1.60)	3.59 (1.50)	5.06 (1.95)	4.21 (1.33)
Control Condition (N = 13)	2.50 (1.49)	1.69 (1.18)	1.46 (0.97)	5.65 (2.02)	2.92 (1.85)	3.46 (1.98) *	3.19 (1.73) *	5.23 (2.09)	4.92 (2.18) *	6.31 (1.55)	5.49 (1.52) *

* Denotes significant main effect.

Results from the one-way ANOVA for task difficulty ratings during the learning phase indicated a significant main effect for the total learning phase, $F(2, 82) = 7.567$, $p = 0.001$. Post hoc comparisons using the Dunnett test showed that the task difficulty rating for the total learning phase for the Control condition was significantly higher than the rating for the ProductWE condition, $p = 0.045$, with a large effect size obtained, $d = 0.98$. The Dunnett Statistical Analysis compares groups with other groups [38] where the result of the homogeneity of variance was significant, $p < 0.005$. The Control condition did not statistically differ from the ProcessWE condition, $p = 0.066$, and the ProcessWE condition did not statistically differ from the ProductWE condition, $p = 0.964$.

Results from the one-way ANOVA for task difficulty ratings during the learning phase revealed no significant main effect for Worked Example 1, $F(2, 82) = 2.545$, $p = 0.085$. However, a significant main effect was found for Worked Example 2, $F(2, 82) = 10.640$, $p = 0.000$. Post hoc comparisons using the Dunnett test indicated that task difficulty ratings for Worked Example 2 were significantly higher in the Control condition compared to both the ProductWE and ProcessWE conditions, with $p = 0.019$ and $p = 0.043$, respectively.

Hypothesis 4 proposed higher perceived task difficulty for participants in the ProcessWE or Control conditions compared to the ProductWE condition during the learning phase. Results showed a significant difference between the Control and ProductWE conditions, but no difference between the Control and ProcessWE conditions. However, for Worked Example 2, the Control condition reported statistically significantly higher task difficulty ratings compared to both the ProcessWE and ProductWE conditions.

Results from the one-way ANOVA for task difficulty ratings during the test phase indicated a significant main effect for the total phase, $F(2, 82) = 4.265$, $p = 0.017$. Post hoc comparisons using the Tukey HSD test showed that the perceived task difficulty rating during the test phase was significantly higher for the Control condition than for both the ProductWE condition, $p = 0.013$ with large effect size, $d = 0.90$, and the ProcessWE condition, $p = 0.049$, medium effect size, $d = 0.73$.

Results from the one-way ANOVA for task difficulty ratings during the test phase showed no significant main effect for Task 1, $F(2, 82) = 2.480$, $p = 0.090$. One-way ANOVA results for task difficulty ratings during the test phase indicated a significant main effect for Task 2, $F(2, 82) = 3.212$, $p = 0.045$. Post hoc comparisons using the Tukey HSD test showed that the task difficulty rating for Task 2 for the Control condition was significantly higher than the task difficulty ratings for the ProductWE condition, $p = 0.041$. Results from the one-way ANOVA for task difficulty ratings during the test phase showed no significant main effect for Task 3, $F(2, 82) = 2.408$, $p = 0.096$.

Hypothesis 6, proposing lower perceived task difficulty for participants in the ProcessWE condition compared to the ProductWE condition during the test phase, did not

demonstrate statistical significance in the experiment, therefore not supporting the hypothesis. Hypothesis 8 states that when learning about the characteristics of substantive communication, participants presented with the ProcessWE condition or ProductWE condition will report lower perceived task difficulty during the test phase than participants presented with the Control condition. Results from the one-way ANOVA for task difficulty ratings indicated a significant main effect for the total task difficulty rating for the test phase. Further, the post hoc comparisons using the Tukey HSD test showed that the task difficulty rating during the test phase for the Control condition was significantly higher than the task difficulty ratings during the test phase for both the ProcessWE condition and ProductWE condition.

In summary, there were five key findings in relation to the aim of the experiment to investigate the participant's cognitive load for the three instructional conditions:

- The ProcessWE condition significantly outperformed the ProductWE condition in both total test performance scores and Task 1 test performance scores. The total test performance score was the sum of three tasks participants completed: a recall task, a near transfer task, and a far transfer task. Each task was scored using clear marking criteria. This aligns with previous findings on the worked example effect, which indicate that process-oriented worked examples enhance cognitive schema construction and automation, resulting in better performance than product-oriented worked examples [13,19,23]. Additionally, integrating principled knowledge in process-oriented worked examples enhances understanding by motivating learners to invest effort into understanding the "why" and "how" [23]. This finding is novel because process-oriented and product-oriented worked examples have primarily been researched in well-structured learning domains such as mathematics, science, and economics [1,5]. While extensive research has demonstrated the effectiveness of worked examples in well-structured learning domains [39,40], there has been less research on their effectiveness in ill-structured learning domains [1,5].
- Mental effort ratings for Worked Example 2 during the learning phase for the ProcessWE and ProductWE conditions were significantly lower than the Control condition. This aspect of CLT warrants further investigation using worked examples in ill-structured learning domains, as the additional information included in these examples can increase cognitive load [40,41].
- Task difficulty ratings during the learning phase for the Control condition were significantly higher than the ProductWE condition. This finding demonstrates that the worked example effect, as predicted by CLT, positively impacted the learning of novice participants, consistent with the effectiveness of worked examples in well-structured learning domains [1,42].
- Task difficulty ratings during the test phase for the Control condition were significantly higher than those for the ProcessWE and ProductWE conditions. This finding aligns with existing CLT research in well-structured learning domains [1,29,43,44].
- Task difficulty ratings for Worked Example 2 during the learning phase for the Control condition were significantly higher than those for both the ProcessWE and ProductWE conditions. This finding is also consistent with existing CLT research in well-structured learning domains [1,29,43,44].

5. Discussion

Research questions were posed to assess test performances, perceived mental effort, and perceived task difficulty among participants in the three instructional conditions. Regarding Research Question 1, comparisons of test performances showed that the ProcessWE condition significantly outperformed the ProductWE condition. However, there was no significant difference between the ProcessWE and Control conditions, despite higher mean scores in the ProcessWE condition. Similarly, no significant difference was found between the ProductWE and Control conditions, contradicting Hypothesis 2. Nevertheless, participants in the ProductWE condition generally outperformed those in the Control condition.

The literature indicates that mental effort and task difficulty ratings are higher for novices who are not presented with worked examples [1,41–44].

Regarding Research Question 2, there were no significant differences in perceived mental effort or perceived task difficulty during the learning phase between the ProcessWE and ProductWE conditions, contradicting Hypotheses 3 and 4. Additionally, no significant difference was found in perceived mental effort ratings during the learning phase between the ProductWE and Control conditions, thus not confirming Hypothesis 3. However, there was a significant difference in perceived task difficulty ratings during the learning phase between the ProductWE and Control conditions, confirming Hypothesis 4, with the Control condition having a higher mean. Furthermore, analyzing individual worked examples revealed significant differences, with Worked Example 2 showing lower perceived mental effort and task difficulty ratings for both ProcessWE and ProductWE compared to the Control condition, confirming Hypothesis 4.

For Research Question 3, investigating perceived mental effort and perceived task difficulty during the test phase between the ProcessWE and ProductWE conditions yielded no significant differences, failing to confirm Hypotheses 5 and 6. Research Question 4 aimed to determine if the ProcessWE or ProductWE conditions led to lower perceived mental effort and task difficulty compared to the Control condition during the test phase. The results showed no significant difference in perceived mental effort between the ProcessWE/ProductWE and Control conditions, contradicting Hypothesis 7. However, a significant difference in perceived task difficulty supported Hypothesis 8. The literature indicates that mental effort and task difficulty ratings are higher for novice learners who are not presented with worked examples compared to novice learners presented with worked examples [1,41–44].

6. Conclusions

This study advances CLT research by offering both theoretical insights and practical applications. Theoretically, this study is one of the first that examined worked examples in an ill-structured learning domain and has shown evidence of the effectiveness of ProcessWEs in ill-structured learning domains, an area that has been largely under-researched [1,5]. Practically, this study contributes to CLT as the instructional materials produced for this study serve as an example of how to structure and format ProcessWEs and ProductWEs within an ill-structured learning domain using video and text annotations. An insight into their presentation and structure is that the text annotations were concise and displayed immediately after the concept was demonstrated in the video. Finally, as the NSW QTM has been found to improve teacher practice and student performance, these materials may inform NSW QTM professional development materials. To further contribute to the field of CLT, more research is required to explore other ill-structured learning domains beyond quality teaching to understand how different instructional conditions support participant engagement and meaning-making. This could take the form of both quantitative and qualitative studies. For example, further experimental studies are required to investigate ProductWEs and ProcessWEs, but also, qualitative studies can provide insight into how participants engage with these worked examples. A promising direction for future research is to examine how novice and expert participants engage with ProcessWEs and ProductWEs in an ill-structured learning domain to better understand their thoughts and processes. This would contribute to CLT research by examining whether there is presence of the expertise reversal effect. The expertise reversal effect asserts that learners require varying levels of support based on their expertise [26,42,45,46], and suggests that the effectiveness of worked examples diminishes as learners become more proficient and build schemas through the continued use of worked examples [42].

Four limitations are identified in this research. Firstly, there were unequal sample sizes for the three conditions, including a small sample size for the Control condition due to participant availability issues. Second, participants all belong to one university. Future studies could involve students from diverse universities. Third, focusing solely

on substantive communication within the NSW QTM may limit result generalization. Fourth, including pauses in Control condition video recordings, while enhancing research robustness, may unintentionally scaffold learning, affecting outcomes.

Overall, this study demonstrated that the use of ProcessWEs in an ill-structured learning domain aided participants in developing relevant schemas for understanding and applying knowledge in the NSW QTM element of substantive communication. Similarly, ProductWEs were found to be more effective than conventional problem-solving methods for learning about substantive communication in the NSW QTM. These findings align with existing CLT research on worked examples in well-structured learning domains for novice participants [19,23,28], as the ProcessWEs, providing principled knowledge and strategic information, improved performance during the test phase in an ill-structured learning domain. However, while there were some significant results aligning with CLT research regarding participants' perceived mental effort and task difficulty ratings, they were inconclusive.

Author Contributions: Conceptualization, G.S., S.A., S.T.-F. and F.P.; methodology, G.S., S.A., S.T.-F. and F.P.; investigation, G.S., S.A. and S.T.-F.; resources, G.S., S.A. and S.T.-F.; data curation, G.S., S.A. and S.T.-F.; writing—original draft preparation, G.S.; writing—review and editing, S.A., S.T.-F. and F.P.; supervision, S.A., S.T.-F. and F.P.. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study was conducted in accordance with the National Statement on Ethical Conduct in Human Research and approved by the Human Research Ethics Committee (HREC) of THE UNIVERSITY OF WOLLONGONG (Ethics Number: 2015/360 and date of approval: 14 August 2018).

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

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the author upon request.

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

References

1. Sweller, J.; Ayres, P.; Kalyuga, S. *Cognitive Load Theory*; Springer: New York, NY, USA, 2011. [CrossRef]
2. Clark, R.C.; Nguyen, F.; Sweller, J. *Efficiency in Learning: Evidence-Based Guidelines to Manage Cognitive Load*; Pfeiffer: San Francisco, CA, USA, 2006.
3. Sweller, J.; van Merriënboer, J.J.G.; Paas, F. Cognitive architecture and instructional design. *Educ. Psychol. Rev.* **1998**, *10*, 251–296. [CrossRef]
4. Sweller, J.; van Merriënboer, J.; Paas, F. Cognitive architecture and instructional design: 20 years later. *Educ. Psychol. Rev.* **2019**, *31*, 261–292. [CrossRef]
5. Kyun, S.; Kalyuga, S.; Sweller, J. The effect of worked examples when learning to write essays in English literature. *J. Exp. Educ.* **2013**, *81*, 385–408. [CrossRef]
6. Simon, H. The structure of ill structured problems. *Artif. Intell.* **1973**, *4*, 181–201. [CrossRef]
7. Cowan, N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav. Brain Sci.* **2001**, *24*, 87–114. [CrossRef] [PubMed]
8. Miller, G.A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* **1956**, *63*, 81–97. [CrossRef]
9. Peterson, L.; Peterson, M. Short-term retention of individual verbal items. *J. Exp. Psychol.* **1959**, *58*, 193–198. [CrossRef]
10. The Education and Training Foundation. The Importance of Cognitive Load Theory. 2024. Available online: https://set.et-foundation.co.uk/resources/the-importance-of-cognitive-load-theory (accessed on 27 June 2023).
11. Chandler, P.; Sweller, J. Cognitive load theory and the format of instruction. *Cogn. Instr.* **1991**, *8*, 293–332. [CrossRef]
12. Eysenck, M.E.; Calvo, M.G. Anxiety and performance: The processing efficiency theory. *Cogn. Emot.* **1992**, *6*, 409–434. [CrossRef]
13. Chen, O.; Paas, F.; Sweller, J. A cognitive load theory approach to defining and measuring task complexity through element interactivity. *Educ. Psychol. Rev.* **2023**, *35*, 63. [CrossRef]
14. Sweller, J. Element interactivity and intrinsic, extraneous, and germane Cognitive load. *Educ. Psychol. Rev.* **2010**, *22*, 123–138. [CrossRef]

15. van Merriënboer, J.; Sweller, J. Cognitive load theory and complex learning: Recent developments and future directions. *Educ. Psychol. Rev.* **2005**, *17*, 147–177. [CrossRef]
16. Sweller, J.; Cooper, G. The use of worked examples as a substitute for problem solving in learning algebra. *Cogn. Instr.* **1985**, *2*, 59–89. [CrossRef]
17. Simon, D.; Simon, H. Individual differences in solving physics problems. In *Children's Thinking: What Develops*; Siegler, R., Ed.; Erlbaum: Hillsdale, NJ, USA, 1978; pp. 325–348.
18. Plass, J.L.; Moreno, R.; Brünken, R. *Cognitive Load Theory*; Cambridge University Press: Cambridge, UK, 2010. [CrossRef]
19. van Gog, T.; Paas, F.; van Merriënboer, J. Effects of studying sequences of process-oriented and product-oriented worked examples on troubleshooting transfer efficiency. *Learn. Instr.* **2008**, *18*, 211–222. [CrossRef]
20. Brooks, C.D. Effects of Process-Oriented and Product-Oriented Worked Examples and Prior Knowledge on Learner Problem Solving and Attitude: A Study in the Domain of Microeconomics. Ph.D. Thesis, The Florida State University, Tallahassee, FL, USA, 2009.
21. Jones, E.C. Cognitive Load Theory and College Composition: Can Worked Examples Help Novice Writers Learn Argumentation? Ph.D. Thesis, Capella University, Minneapolis, MN, USA, 2014.
22. Wong, R.M.; Adesope, O.O.; Carbonneau, K.J. Process- and product-oriented worked examples and self-explanations to improve learning performance. *J. STEM Educ. Innov. Res.* **2019**, *20*, 24–31.
23. Ohlsson, S.; Rees, E. The function of conceptual understanding in the learning of arithmetic procedures. *Cogn. Instr.* **1991**, *8*, 103–179. [CrossRef]
24. New South Wales Department of Education and Training. Professional Support and Curriculum Directorate. In *Quality Teaching in NSW Public Schools: Discussion Paper*; Department of Education and Training: Sydney, NSW, Australia, 2003.
25. New South Wales Department of Education and Training. Learning and Leadership Development Directorate. In *Quality Teaching in NSW Public Schools. A Classroom Practice Guide*; Department of Education and Training: Sydney, NSW, Australia, 2006.
26. Research Centre of Teachers and Teaching. Quality Teaching Rounds. Available online: https://www.newcastle.edu.au/research/centre/teachers-and-teaching/quality-teaching-rounds (accessed on 28 April 2024).
27. Owens, K. Teachers' development of substantive communication about mathematics. In *MERGA 28: Building Connections: Research, Theory, and Practice*; MERGA Inc.: Melbourne, VIC, Australia, 2005; Volume 2, pp. 601–608.
28. Chen, O.; Retnowati, E.; Chan, B.K.Y.; Kalyuga, S. The effect of worked examples on learning solution steps and knowledge transfer. *Educ. Psychol.* **2023**, *43*, 914–928. [CrossRef]
29. Gupta, U.; Zheng, R.Z. Cognitive load in solving mathematics problems: Validating the role of motivation and the interaction among prior knowledge, worked examples, and task difficulty. *Eur. J. STEM Educ.* **2020**, *5*, 5. [CrossRef]
30. Paas, F. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *J. Educ. Psychol.* **1992**, *84*, 429–434. [CrossRef]
31. Schmeck, A.; Opfermann, M.; van Gog, T.; Paas, F.; Leutner, D. Measuring cognitive load with subjective rating scales during problem solving: Differences between immediate and delayed ratings. *Instr. Sci.* **2015**, *43*, 93–114. [CrossRef]
32. Marcus, N.; Cooper, M.; Sweller, J. Understanding instructions. *J. Educ. Psychol.* **1996**, *88*, 49–63. [CrossRef]
33. Hsu, C.; Eastwood, J.; Toplak, M. Differences in perceived mental effort required and discomfort during a working memory task between individuals at-risk and not at-risk for ADHD. *Front. Psychol.* **2017**, *8*, 407. [CrossRef] [PubMed]
34. New South Wales Department of Education and Training, Professional Support and Curriculum Directorate. Quality Teaching in NSW Public Schools [Kit]: Continuing the Discussion about Classroom Practice. Available online: https://catalogue.nla.gov.au/catalog/3069717 (accessed on 28 April 2024).
35. Quality Teaching in New South Wales Public Schools. Professional Support and Curriculum Directorate. In *Continuing the Discussion about Classroom Practice*; Department of Education and Training: Sydney, NSW, Australia, 2003.
36. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; L. Erlbaum Associates: Hillsdale, NJ, USA, 1988.
37. Allen, M. *The Sage Encyclopedia of Communication Research Methods*; SAGE Publications, Inc.: Thousand Oaks, CA, USA, 2017; Volume 1–4. [CrossRef]
38. Ruxton, G.; Beauchamp, G. Time for some a priori thinking about post hoc testing. *Behav. Ecol.* **2008**, *19*, 690–693. [CrossRef]
39. Atkinson, R.; Derry, S.; Renkl, A.; Wortham, D. Learning from examples: Instructional principles from worked examples research. *Rev. Educ. Res.* **2000**, *70*, 181–214. [CrossRef]
40. Renkl, A. The worked-out examples principle in multimedia learning. In *The Cambridge Handbook of Multimedia Learning*; Mayer, R.E., Ed.; Cambridge University Press: New York, NY, USA, 2005; pp. 229–245. [CrossRef]
41. Puma, S.; Tricot, A. Cognitive load theory and working memory models. In *Advances in Cognitive Load Theory: Rethinking Teaching*; Tindall-Ford, S., Agostinho, S., Sweller, J., Eds.; Routledge: London, UK; Taylor & Francis Group: Abingdon, UK, 2020; pp. 30–40. [CrossRef]
42. Kalyuga, S.; Sweller, J. Cognitive load and expertise reversal. In *The Cambridge Handbook of Expertise and Expert Performance*; Ericsson, K., Hoffman, R., Kozbelt, A., Williams, A., Eds.; Cambridge Handbooks in Psychology; Cambridge University Press: Cambridge, UK, 2018; pp. 793–811. [CrossRef]
43. Zhang, L.; Yu, Y.; Liu, Y.; Qu, K. Influence of material difficulty and strategy knowledge on worked example: Learning strategy selection for solving probability problems. *Curr. Psychol.* **2023**, *42*, 26477–26490. [CrossRef]

44. Overson, C.E.; Hakala, C.M.; Kordonowy, L.L.; Benassi, V.A. (Eds.) *In Their Own Words: What Scholars and Teachers Want You to Know about Why and How to Apply the Science of Learning in Your Academic Setting*; Society for the Teaching of Psychology: Washington, DC, USA, 2023.
45. Lovell, O. *Sweller's Cognitive Load Theory in Action*; John Catt Educational Ltd.: Woodbridge, UK, 2020.
46. Chen, O.; Kalyuga, S.; Sweller, J. The expertise reversal effect is a variant of the more general element interactivity effect. *Educ. Psychol. Rev.* **2016**, *29*, 393–405. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.