Effects of Worked Examples on Stress, Cognitive Loads, and Performance in Online Collaboration

オンライン協働におけるワークトエグザンプルの
ストレス、認知負荷、パフォーマンスへの影響

A Dissertation Presented to
the Division of Education,
the Graduate School of International Christian University,
for the Degree of Doctor of Philosophy

国際基督教大学　大学院
教育学研究科提出博士論文

Kudo, Masayuki
工藤　雅之

November 21, 2013
2013 年 11 月 21 日

Acknowledgements

I would like to express my deepest appreciation to Professor Insung Jung for her tremendous encouragement and guidance throughout these six years. Without her persistent help and enthusiastic guidance, this dissertation would not have been possible. Her patience is only surpassed by my shortcomings as a scholar. Her willingness to go above and beyond the call of duty to mentor me via ICT has provided me with professional support I hope some day to duplicate. It will be a rewarding challenge for me to provide my students with the same kind of mentorship I received.

I can hardly thank both Professor Teruyoshi Sasaki and Professor Toshiaki Sasao enough, for their support and cooperation as members of my committee. Professor Sasaki was generous with his encouragement. Professor Sasao was there by my side with feedback, even when he was thousands of miles away in Poland. I wish nothing but good fortune and happiness for both professors and their families. I also thank Professor Hidenori Fujita, who provided me with advice during the initial stage of my project. Without his advice my study would not have gotten off the ground.

I would like to thank Dr. Isao Watanabe who has shown me what it means to be an ICU doctoral student. I appreciated the countless advice as both Senpai and colleague. Also, I would like to thank Dr. Kota Tamai who stood by me during the final stages of this dissertation. Without his precise academic contributions, this project would not have come to fruition. I share the credit of my work with Mr. Tsuchida for managing experiments and data analysis. He has been more than I could ask of a colleague, or a friend. I would like to show my appreciation for Mr. Joseph Tomasine, who supported me first with English writing, and with his encouragement when I was downtrodden.

I would like to thank all faculty members of the department of Media Communication Design of Hokkaido Institute of Technology, lead by Dean Suzuki for supporting me from the start of this long process. Without your understanding, I could surely not have come this far. In addition, I would like to thank Dr. Ferdinand Pitagan, Dr. Yayoi Anzai, Ms. Sookkyoung Choi, Ms. Aya Fukuda, and Mr. Gibran Garcia for their cooperation as colleagues at ICU. The conversations we had during my visits to ICU contributed more than I can say.

I cannot forget to express my appreciation for my students, from my seminar groups in recent years and of course those who took part in the experiments. I hope you have learned as much from me as I have from you. This dissertation represents but a part of the knowledge and learning made possible by your contributions.

I would express my sincere appreciation to Gando-san, Koyanagi-san, and Osaki-san, and other staff members of ICU. No one could have better understood the special needs of a student like me, working and living in Sapporo. The support they provided cleared obstacles that would otherwise have spelled disaster for this dissertation.

Finally, I would like to express my deep gratitude to my parents. Schooling for them stopped with a high school diploma, obtained through night classes. Needless to say, this modest academic background did not stop them from encouraging me to achieve academically. This became one of my greatest sources of motivation. I am grateful for their unconditional belief in me and for their endless support.

Last but not least, let me express my gratitude to my beloved wife, Hiromi who spent sleepless nights with and was always my support in the moments when there was no one to answer my queries. She has held my hand for almost 20 years, since this adventure began. This dream of ours would not have come true without her.

Abstract

Many agree online collaboration is a promising educational intervention, realizing interactions among people with diverse learning backgrounds for the 21st century. Online collaboration is known to be beneficial in facilitating the exchange of ideas, opinions, facts, experiences and expectations (McConnell, 2000), and supporting learning outcomes for varied learner demographics, regardless of gender, major, grade, or race (Barkley, Cross, & Major, 2005).

In spite of the advantages of online collaboration, prior studies have reported less-favorable or mixed results. In particular, a study by Jung, Kudo, and Choi (2012) with English as a foreign language (EFL) students in Japan has revealed that online collaboration increases participants' stress level due to four types of stressors: Type I – lack of confidence or self-efficacy in language related task competencies, Type II – poor instructional design for online collaborative work, Type III – problems related to technology use, and Type IV – the online interactions with others required by the collaboration process. In order to mitigate unnecessary stress caused by these stressors, previous research has indicated the need to develop instructional strategies tailored to online collaborative learning. Unfortunately, very few studies have empirically tested strategies for, and approaches to, the neutralization of stressors in online collaboration. Among those few, Jung et al. (in press) and Kudo (2013) have recommended the use of worked examples (WEs) as theorized and developed within cognitive load theory.

The present study investigated the impact of two types of WEs (WE with and without prompts) on four types of stressors, three kinds of cognitive loads (intrinsic, extraneous and germane) in relation to levels of language self-efficacy (LSE) during online collaborative learning. Collaborative composition performance was also

examined by the alteration of two instructional strategies. Three research questions posed include:

1) What type of worked examples would be effective in reducing stressors of EFL learners with two different levels of language self-efficacy in an online collaborative learning context;
2) What type of worked examples would be effective in reducing cognitive loads of EFL learners with two different levels of language self-efficacy in an online collaborative learning context;
3) What type of worked examples would be effective in contributing better collaborative performance of EFL learners in an online collaborative learning context?

In order to answer these questions, two quasi-experiments were carried out at a local institute of technology, employing 60 junior students for Study 1 and 80 freshmen and sophomore students for Study 2. The participants of both Study 1 and 2 participated in the six-week online collaborative composition project and were grouped into triads in Study 1 and dyads in Study 2. Study 1 compared the effects of two instructional strategies on stressors, on cognitive loads, before and after the treatment in relation to the levels of LSE. In addition, collaborative composition performance was examined by comparing the difference of instructional strategies. Study 2 followed the same experimental procedure as Study 1, except it employed a control group that utilized no instructional strategy and the stressor scale was measured once at the end of experiment. Stressor scale (Jung et al., 2011) was used for the measurement of Type I to IV stressors and Foreign language self-efficacy scale (Shaw, 2006) was used to determine task as well as skill confidence in foreign language usage. The mean values of stressors as well as cognitive loads as dependent variables were compared across the groups. *ESL*

composition profile (Jacobs, Zinkgraf, Wormuth, Hartfiel, & Hughey, 1981) was employed for the evaluation of collaborative composition performance.

In addressing Research Question one, a three-way analysis of variance (ANOVA) F-tests examined each factor of treatments (WE with and without prompts), LSE (high and low) and Before-After on the mean values of four types of stressors. The tests found the interaction effect between Before-After and treatments in the low LSE group. The result revealed Type I stressor to be reduced by the use of the WE with prompts for the low LSE group in the After condition. This finding added a new characteristic – stress mitigation effect – to worked examples and supported the arguments made by Niculescu et al. (2009) that cognitive demand, which causes stress, can be conceptualized together with mental effort. The interaction effect found in Study 1 also revealed the possibility that LSE determined the effect of additional prompts. This suggested that prompts played a pivotal role in the mitigation of Type I stressor for low LSE learners, functioning as scaffolding. Prompts must have become a "more able other" (p.15) as argued in Davis (2003) to assist the low LSE learners in utilizing the WEs for their collaborative work. However, Study 2 did not replicate these results on Type I stressor and appeared contradicting to the findings from Study 1.

Study 1 also revealed a primary interaction between Before-After and treatments on Type IV stressor. The mean value of Type IV stressor of WE without prompts group showed a higher degree of Type IV stressor in the After condition. The contrast with Type I stressor suggested that the group utilizing WE without prompts experienced difficulties in communication during online collaboration as a result of the lack of prompts. The results of Study 2 on Type IV stressors did not replicate the result of Study 1, either. The inconsistency was explained by the difference in LSE level and the readiness of the participants for the collaborative task. Neither Type II nor III stressors in Study 1 and 2 were influenced by treatments or LSE. This was accounted by

Type II and III stressors being out of the range of scaffolding that the WE provided. LSE did not cover the confidence of the two respective areas.

Regarding Research Question two, two-way ANOVA F-tests were used to examine the difference in three kinds of cognitive loads (intrinsic, extraneous and germane) and the effects of two between-subjects factors, treatments (WE with and without prompts) and the LSE (high and low). Extraneous and germane cognitive load did not show the interaction in both experiments. The high LSE group revealed a lower degree of intrinsic cognitive load than the low LSE group when they utilized the WE with prompts. This confirmed the effectiveness of prompts in maintaining low intrinsic cognitive load for the high LSE group. The result was in accordance with the previous findings that the higher the self-efficacy one obtains, the larger the working memory capacity (Hoffman & Schraw, 2009) and that prompts could increase intrinsic cognitive load for such individuals (Berthold & Renkl, 2009; Berthold et al., 2011). In Study 2, the main effect of LSE only influenced the perceptions of intrinsic and extraneous cognitive loads.

On germane cognitive load, the high LSE group indicated a higher score than the low LSE group as a result of the simple main effect of LSE. The result suggested that the cognitive strategies of the high LSE group could have been enhanced. However, treatments did not affect the germane cognitive load, indicating that prompts did not help the participants to generate more cognitive strategies. This result seemed to contradict the result on Type I stressor reduction. The contradiction calls for future research.

In addressing Research Question three, t-tests on two treatments compared the differences in collaborative composition performance in Study 1 and found no differences. A one-way ANOVA F-test in Study 2 comparing three groups including a control group detected Content and Organization to be influenced by the treatments.

The Content scores of the two treatment groups were higher than the control group, which suggested that WEs, whether with or without prompts, helped improve the quality of the Content. Organization was affected by the use of WE with prompts, indicating that prompts added to WE were the source of improvement for compositional structure. The result implied that prompts directed the participants' attention to the structure of the example, allowing them to carefully model off the WE and produce better compositions.

Nevertheless, Study 2 failed to replicate the results of Study 1 and indicated some inconsistent results. For instance, the mean scores of Type I stressor observed in Study 1 and 2 produced opposing data. While the low LSE group showed a difference between treatments in Study 1, Study 2 revealed that the high LSE group was affected. Also, Study 1 showed a significant difference between the treatments, Study 2 found a difference between the control group and the treatment groups, consequently failing to show the difference in the treatments. The level of LSE and readiness for the task may provide explanations for the inconsistency. In fact, comparing the mean scores of LSE between the two experiments, the score was significantly higher in the participants of Study 1. Future empirical studies are required to explain the inconsistency.

In conclusion, despite limitations, the results of Study 1 and 2 stated several implications and suggestions for online collaboration dealing with the EFL learners of low LSE. WE with prompts were confirmed to reduce the influence of Type I stressor, especially for the low LSE group. The study also confirmed that EFL learners of low LSE were able to contribute to the collaborative composition via the help of this instructional strategy. It proposed the use of additional scaffolding with prompts to expand the effectiveness of WE for learners of low LSE.

The results of both Study 1 and 2 suggested that LSE was influential in developing and facilitating online collaboration using foreign language. The finding

suggested that LSE was effective not only for the perception of Type I stressor, but also cognitive loads. In online collaboration, where learners work on their own and collaborate with peers, self-efficacy directly affected their participation and performance. It was found to be a powerful instructional strategy that made it possible for EFL learners to borrow both structural features and content from the example. In search of appropriate instructional strategy, the use of double-content WEs is recommended for EFL collaborative composition administered online. Combined with self-explanation prompts, double-content WEs were made accessible to learners of low LSE.

Finally, the study confirmed the importance of scaffolding. The WE contributed to composition completion in EFL by scaffolding learners of low LSE. Prompts were found to make a strong impact as scaffolding on leaners. It is essential to scaffold learners' lack of confidence with appropriate support to fill the gap in knowledge or skills, in order to realize effective and enjoyable online collaborative learning.

論文要旨

21 世紀のもっとも有望な教育手法として、多くの研究者がオンラインによる協働学習に注目している。なぜなら、オンライン協働学習はさまざまな学習背景を持つ人々の相互交流を可能にするからである。整備されたオンライン協働学習環境は、アイディア、意見、事実、経験や期待など、学習者同士のやりとりを促進する利点があり (McConnell, 2000)、学習者の性別、専攻、学年や人種などの個人要因に関わりなく有効である (Barkley, Cross, & Major, 2005) ことが広く知られている。

このように、オンライン協働学習には多くの利点があるにも関わらず、先行研究の一部においては、効果が少ないものや利点を強調できないという結果も報告されている。Jung, Kudo, and Choi (2012) の日本における第二言語としての英語利用環境下での研究では、オンライン協働学習に従事する場合、ストレスを発生させる４つの要因があることが報告された。それらは、（１）課題達成での言語に関わる自信、あるいは自己効力の欠如、（２）オンライン協働の為の不十分な教授設計、（３）技術利用に関する諸問題、そして （４）協働過程に必要なオンラインでの相互交流、である。

不必要なストレスを緩和する為に、先行研究ではオンライン協働学習をサポートするための教授方略の必要性が議論されている。しかしながら、オンライン協働学習でのストレス要因に応対する研究は数少ない。まして具体的な方略、取り組みによってオンライン協働学習のストレス要因を緩和する実証的な研究に至っては非常に少ないといえる。数少ない研究の中で、Jung et al. (in press) や Kudo (2013) は、認知負荷理論が提唱するワークトエグザンプル (worked example) の利用を提案している。ワークトエグザンプルとは、どのように問題や課題を解決に導くか段階的な模範例や手法を提示する教授方略である。

本研究は、オンライン協働学習において、２つの形式のワークトエグザンプル（プロンプト有り、無し）が、言語自己効力（language self-efficacy = LSE）のレベルとの関わりにおいて、４つのストレス要因（タイプ I, II, III, IV）と３つの認知負荷 (cognitive loads)、すなわち内在的 (intrinsic) 認知負荷、外在的 (extraneous) 認知負荷、妥当な (germane) 認知負荷に与える効果、および作文課題に与える効果を明らかにする事を目的とした。研究課題として挙げたのは以下の３つである。

１） どの形式のワークトエグザンプルが、オンライン協働学習で２つの LSE レベルの EFL 学習者のストレス要因緩和に対して効果的か？

２） どの形式のワークトエグザンプルが、オンライン協働学習で２つの LSE レベルをもつ EFL 学習者の認知負荷軽減に対して効果的か？

３） どの形式のワークトエグザンプルが、オンライン協働学習で EFL 学習者の協働での作文成績に対して効果的か？

上記の研究課題に答える為に、本研究では地方にある工業大学で２つの準実験を行い、実験１に 60 名（3 年生)、実験２では 80 名（1 年生と 2 年生）が参加した。参加者は、実験１では３人一組、実験２では２人一組で６週間のオンライン協働作文課題に従事した。実験１は、２つの教授方略の違いによりストレス要因と認知負荷が処遇の前後において、自己効力の高さとの関連でどう変化するかを比較した。さらに、協働による作文は、教授方略の違いによる差で検証された。実験２では、実験１と同じ実験手順が踏襲されたが、まったく何の教授方略も付与しない統制群を置き、ストレス要因調査を二度から一度に変更した。Type I から IV のストレス要因の計測には Stressor scale (Jung et al., 2011) が使用され、外国語でのタスクとスキルの自信の計測には、Foreign language self-efficacy scale (Shaw, 2006) が使用された。協働作文を評価するにあたっては、*ESL composition profile* (Jacobs et al., 1981)を使用した。

研究課題１に関しては、４つの型のストレス要因の平均値に対し、三要因の分散分析を処遇要因、LSE 要因、そして前後要因に対して行った。その結果、低 LSE 群において、前後と処遇に交互作用が認められた。低 LSE 群では、プロンプト付ワークトエグザンプルを用いると、処遇の後の状態で Type I ストレス要因を軽減できることが判った。このことは、ワークトエグザンプルに「ストレス緩和効果」と言うべき新たな知見を加え、Niculescu ら (2009) のストレス反応を引き起こす認知負担は、心的努力（mental effort）と要素を共有するという説をサポートするものである。また実験１で観察された交互作用により、LSE がプロンプトの効果を左右したことも認められた。この結果は、低 LSE 学習者にとってプロンプトが、ストレス軽減に対し足場掛けとして機能し、枢軸的な効果があったことを示している。Davis (2003) が言うようにプロンプトは「有能な他人」として機能し、協働作業を行う上で必要なワークトエグザンプルが使用できるよう低 LSE 学習者をほう助したことを示唆している。しかしながら、実験２は Type I ストレス要因に関する実験を再現できず、実験１と相対する結果を示した。

実験１では Type IV ストレス要因に対して、前後と処遇の交互作用を認めた。プロンプト無しのワークトエグザンプルグループの Type IV ストレス要因の平均値が、処遇の後において高かった。Type I ストレス要因との対照で考えると、プロンプト無しのワークトエグザンプルを使ったグループは、ワークトエグザンプルの使用をほう助するプロンプトが提示されなかったことが原因で、ワークトエグザンプルを利用する事ができず、オンライン協働作業での意思疎通で苦労したことが伺える。実験２での Type IV の結果も、実験１の結果を再現する事ができなかった。この不一致は、LSE の差や協働課題へのレディネスが原因と考えられる。両実験で Type II、Type III ストレス要因は処遇と LSE の効果に影響されず、本研究で使用したワークトエグザンプルが Type II と III の影響する範囲では「足場掛け」効果が充当せず、LSE がこの２つの分野に関わりが薄

い事を示す。

研究課題２について、二要因の分散分析を３つの認知負荷（内在的認知負荷、外在的認知負荷、妥当な認知負荷）の平均値の違いに対して被験者間要因である処遇と LSE で行った。分析の結果、実験１では処遇と LSE の交互作用が認められたが、外在的な認知負荷と妥当な認知負荷では交互作用を認めなかった。プロンプト付きワークトエグザンプルを使用すると高 LSE 群では、低 LSE 群に比べて内在的な認知負荷の値が低く観測された。これは高 LSE 群にとってプロンプトが、内在的認知負荷を低く抑える効果を確認したものである。この結果は、高い自己効力を持つものほど大きな作動記憶を持つ (Hoffman & Schraw, 2009) という知見や、プロンプトは内在的な認知負荷を押し上げる (Berthold & Renkl, 2009; Berthold et al., 2011) という知見と一致するものである。実験２では、内在的認知負荷、外在的認知負荷の両方とも LSE 要因の主効果しか観測できなかった。

実験１では、妥当な認知負荷の値が低 LSE 群と比較すると高 LSE 群で高かった。この結果は、低 LSE 群と比較して高 LSE 群でより多くの認知方略が醸成されたことを示している。しかしながら、本来差が出るはずの処遇における差は見られなかった事から、プロンプトの有無が認知方略の効果的な醸成に寄与するとは言えないことから、この結果は Type I ストレス要因軽減の結果に矛盾する。今後の研究課題としたい。

研究課題３に関して、実験１は２つの処遇において、それぞれの協働作文評価項目に照らして得点の平均値に t-test を行ったが、違いは認められなかった。実験２においては、統制群を含んだ３グループにおける評価項目の平均値に対して一要因の分散分析を行った結果、Content と Organization の項目で２つのワークトエグザンプルを使用したグループは、統制群に比べて高い値を示し、２つの処遇の効果が確認された。これは、プロンプトの有無にかかわらず、ワークトエグザンプルが Content 項目内容を高めることを示している。Organization

は、プロンプト付きのワークトエグザンプルに効果が見られ、プロンプト付きワークトエグザンプルが作文構成の改善の源泉になっていると考えられる。この結果は、プロンプトが参加者の注意を例文の Organization に向けることで、参加者がより良い作文を目指してワークトエグザンプルを注意深く模範として利用したことを示す。

しかしながら、実験１と２には、いくつかの不一致が見られた。たとえば、実験２は実験１を再現する事ができず、LSE の高低で逆の反応が観測された。Type I ストレス要因については高 LSE 群が処遇の差の効果を受けたのに対して、実験２では、低 LSE 群が処遇の差の効果を受けた。また、実験１は処遇の差を観測したにもかかわらず、実験２では統制群との差は確認できたものの、２つのワークトエグザンプルの間には差は確認できなかった。実験１では、低 LSE 群が影響を受けたが、実験２では高 LSE が影響を受けた。この差は LSE の差と就学年の違いによる課題に対する学習者のレディネスの差によるものであると考えられ、実際に２つの実験での被験者の平均の LSE の平均値を比べると、実験１の被験者が統計的に有意に高かった。より深い知見を得る為には、LSE とワークトエグザンプルの関係をより具体的に検証する研究が望まれる。

結論として、この研究には制限があるものの、実験１、２の結果から低 LSE を有する EFL 学習者をオンライン協働学習に含むためのいくつかの示唆と提言を得る事ができた。プロンプト付きのワークトエグザンプルは、特に低 LSE 学習者の Type I ストレス要因を減ずる効果がある事が確認され、低 LSE 学習者も学習方略の助けを借りて協働作文課題に貢献し、完遂できることを確認した。その際には、低 LSE 学習者に対してワークトエグザンプルの効果を拡張する為に「足場掛け」を付与し、学習者をサポートすべき事が提案された。

本研究の各実験から得られた知見は、オンライン協働環境を開発、実行する上で、LSE が深い影響力を持つ事を示しており、この知見は LSE が Type I ストレス要因だけでなく認知負荷にも影響を与えることが判った。学習者自身が協

働他者と自らの責任で協働しなければならないオンライン協働環境では、自己効力が参加のみならず、学習成果にも直接的に影響することを示している。またオンライン協働での適切な学習方略設定を目指すにあたり、EFL 環境下での協働作文課題については“２重コンテンツのワークトエグザンプル”の使用が提案された。２重コンテンツのワークトエグザンプルは、EFL 学習者が手本から文章の構造や内容までも「借りる」ことのできる効果的な教授方略であり、自己説明プロンプトと共に利用することで、低 LSE 学習者もこのスタイルのワークトエグザンプルからの恩恵を享受できる事が確認された。

最後に、本研究では足場かけの重要性を確認した。ワークトエグザンプルは、英語能力の欠如を補い「足場をかける」事で低 LSE 学習者の作文課題の遂行に貢献した。プロンプトもまた低 LSE 学習者に対して強固な足場かけ効果がみとめられた。知識やスキルの欠如を補い、学習者の自信をしっかりした手段でサポートすることは、効果的で楽しさを実感できるオンライン協働学習環境を整えるうえで、極めて重要である。

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

Background of the Study

As several scholars and educators (e.g., Barkley, Cross, & Major, 2005; Dillenbourg & Schneider, 1995; Johnson & Johnson, 1999) have claimed, collaboration can be a means to promote social interactions in which learners from diverse learning backgrounds work interactively on a joint solution to a problem. In a well-designed collaborative learning environment, even less-competent learners engage in learning tasks by being encouraged to take ownership of their learning process. This consequently exempts instructors from being merely a vehicle for information transmission (Sawyer, 2006). Moreover, when learners are encouraged to explore various options to solve problems they tend to actively engage with peers (Alavi, 1994; Kirschner, 2001) and achieve greater results compared with peers who adopt a more individual approach to their studies (Johnson & Johnson, 2004).

With the proliferation of Information and Communications Technologies (ICTs), collaborative learning has become the "heart and soul" (Palloff & Pratt, 2005, p.6) of online coursework. ICTs allow more participants to go beyond the constraints of time and location, realizing a liberal educational environment. Online collaboration is more democratic than face to face collaborative work because it makes participation possible for those who are shy, less verbally articulate, more considered, and slower (Clark, 2003). Recent research has revealed that online collaboration improves the richness and quality of the learning experience (Bernard, Rojo de Rubalcava, & St-Pierre, 2000), by facilitating the exchange of ideas, opinions, facts, experiences and expectations (McConnell, 2000). Furthermore, online collaboration ensures the

effectiveness of individual learning across demographics such as sex, majors, grades, or race (Barkley et al., 2005). The collaborative learning environment can even promote and assist the learning of those with diverse experiences and backgrounds by integrating less competent students with more advanced peers in the same learning community (Gerlach, 1994).

Despite the advantages of online collaboration, not all research findings are positive. Simpson (2000) has pointed out that the distance-learning environment, which utilizes online collaboration, generates stressful learning conditions. Gunawardena (1995) has highlighted the issue of social presence among graduate school students engaging in collaborative learning tasks, discovering that the lack of social context cues inherent in computer mediated communications (CMC) can cause negative reactions. In addition to the social issues, Mäkitalo, Weinberger, Häkkinen, Järvelä, and Fischer (2005) point out that ineffective online collaboration results from a lack of clear structure and detailed plans for online discussion. Issues arising from the choice of learning tasks have also been identified, namely task complexity (Kirschner, Paas, Kirschner, & Janssen, 2011). As these studies show, there is a serious need for appropriate instructional strategies and designs for mitigating unnecessary stress, which has been tailored specifically for the online collaboration. Without such strategies and designs, the online learning environment is prone to multiple sources of barriers that prevent learners from learning effectively and efficiently (Jung, Kudo, & Choi, in press).

Lazarus and Folkman (1984) have argued that stress is a reaction to a task or event perceived by the learner to be beyond her capabilities; thus, what may be stressful for one learner is not necessarily stressful for another. A good example of this is how a positive cognitive appraisal of stressors can actually become a source of positive emotion, motivation, and better learning outcome, despite the variety of undesirable

effects of stress that have been documented (e.g., Folkman & Moskovitz, 2000; Lazarus, 1990; Palmer, Cooper, & Thomas, 2003). Bandura (1982) has theorized that when one has a strong sense of self-efficacy, she is more likely to persist in her efforts until she succeeds. Bandura (1993) further explains that the strength of self-efficacy is a good predictor for behavior change, and has hypothesized that self-efficacy influences the level of stress experienced by the learner. When one has the tools to overcome stress, he argues, one expends more energy and persists longer to turn the stress into a positive outcome.

Bandura (1995) adds that when one identifies a source of stress to be beyond their coping capability, external pressure brings about a stress reaction. In contrast, when one identifies the task to be within their capacity, the sources do not make her distressed. Learners with high self-efficacy are thus more tolerant to sources of stress as compared to the lower efficacious learners. However, little empirical research has explored the relationship between stress and self-efficacy, specifically in regards to the online collaborative environment.

Furthermore, little research has been conducted on the sources of stress under online collaborative learning condition until recently. There have, however, been some developments in this area. Technical problems, collaboration itself, and the lack of trust among learners have been identifies as crucial stress factors in online collaboration, more so than in traditional face-to-face classrooms (e.g., Cohen & Gibson, 2003; Salmon, Allan, & Giles, 2000). A growing number of studies have investigated sources of stress (Al-Fudail & Mellar, 2008; Lawless & Allan, 2004; Allan & Lawless), in an attempt to clarify stress factors in order to counteract stress using instructional strategies.

A recent study carried out by Jung, Kudo, and Choi (2012) has identified four types of stressors in online collaboration commonly found in the particular learning

condition of 226 English as a foreign language (EFL) college students in Japan: (1) lack of confidence or self-efficacy in language related task competencies, (2) poor instructional design for online collaborative work, (3) problems related to technology use, and (4) the online interactions with others required by the collaboration process. The four categories of stressors have turned out to be the index for developing instructional strategies for stressor reduction. Studies such as Jung et al. have highlighted the need to develop appropriate instructional strategies tailored specifically to online collaborative learning in order to mitigate against the unnecessary influence of stressors.

Few studies have pursued practical instructional strategies for improving online collaboration by engaging the issue of establishing countermeasures for stress. Even fewer studies have experimented with specific strategies for the prevention or neutralization of stressors. Jung et al. (in press) recommend eight instructional strategies for effective and less stressful online collaboration based on their previous work (Jung, Kudo, & Choi, 2011), cataloging instructional design strategies. One of the eight instructional strategies uses worked examples to eliminate a category of stressors. Worked examples in this work follow the model theorized and developed by cognitive load theorists, namely Sweller and Cooper (1985).

The body of cognitive load theory (CLT) research agrees that worked examples are proven to reduce cognitive load and promote effective learning by scaffolding the learners' process to task completion (Dennen, 2004; Jonassen, 1999; van Merriënboer, 1997). Sweller (2006) points out that worked examples can eliminate the risk of students randomly applying their limited cognitive resources. Learners increase their probability of success and reduce the need for trial and error or means-ends analysis by using the example as a prototype and adhering to the suggested solutions.

Another benefit of the worked example is that it reduces the extraneous load

that does not contribute to learning, in effect moderating cognitive demand (Renkl et al., 2009; van Gog, Paas, & van Merriënboer, 2006). A body of research agrees that when leaners concentrate on applying the principles presented in the worked example, they avoid wasting working memory for problem solving (Kalyuga, Ayres, Chandler, & Sweller, 2003; Renkl, 2002; Spanjers, Wouters, van Gog, & van Merriënboer, 2011; van Gog & Rummel, 2010).

Worked examples are also known to be particularly helpful for initial cognitive skill acquisition (Renkl, Hilbert, & Schworm, 2009; Van Gog, Kester, & Paas, 2011). Novices or learners with less or limited pre-existing knowledge benefit (Atkinson, Derry, Renkl, & Wortham, 2000; van Gog & Rummel, 2010) because they are assumed to consume less cognitive resources if given steps to the solution (Kalyuga, Ayres, Chandler, & Sweller, 2003; Renkl, 2002; Spanjers, Wouters, van Gog, & van Merriënboer, 2011).

Recent work by Hübner, Nückles, and Renkl (2010) has used worked examples to make up for the proficiency gap. Worked example successfully support German high school students who lacked sufficient knowledge to compose learning journals. Diao and Sweller (2007) also confirm the improvement of novice EFL learners' reading comprehension via worked examples. Observation and borrowing (Sweller, 2006) are the major benefits argued in the recent research on worked examples. Despite obvious benefits, little research has focused on the effectiveness of reducing the influence of stressors via the use of worked examples. This is especially true of the task related stressor. There is merit in investigating the possibility of stress mitigation via worked examples, given the benefits of observation and borrowing.

Further literature on worked examples has discovered that adding prompts – questions or short directions induces more effective self-explanation, elicitations of problem states (Schworm & Renkl, 2007) and learning strategies (Hübner et al., 2010),

leading to a better, deeper, and richer set of learning outcomes (Berthold, Nückles, & Renkl, 2007; Schworm & Renkl, 2007), in addition to fostering transfer ability (Renkl, 2005). Prompting is a technique used to induce self-explanations through questions or elicitation. Chi, Bassok, Lewis, Reimann, and Glaser (1989) have confirmed that prompts foster learning strategies that support deeper learning and consequently lead to a more effective learning experience (e.g., Renkl, 2005, Rummel, Spada, & Hauser, 2006; Schworm & Renkl, 2007).

Amidst the research in favor of worked examples with prompts, Hilbert, Renkl, Schworm, Kessler, and Reiss (2008) and Renkl, Hilbert, and Schworm (2009) have reported contradictory findings. They have argued that prompting could in fact increase cognitive load, potentially causing overload for learners with less or limited pre-existing knowledge. They recommend more empirical studies be done on the effectiveness of worked examples with or without prompts in relation to the maintenance of cognitive load, especially for those who lack fundamental knowledge of the task. It is plausible that the combination of worked examples with prompts could successfully scaffold the learners' knowledge sufficiently to contain cognitive load and prevent stress. This would, however, be limited by the learners' level of self-efficacy.

While many effects of worked examples are documented, there seems to be a very few studies using worked examples with and without prompts to mitigate the influence of stressors. The investigation on the relationships between worked examples with prompts, cognitive load and self-efficacy leads us to discover a gap in the research. Furthermore, the research to date on online collaboration leaves out the concept of self-efficacy, despite its influence on human action, including stress reactions. In addition, current research has not fully investigated this combination of elements with regards to learners of low self-efficacy or lower achieving learners working online.

The goal of the present study is to fill this gap in the literature. The experiment focuses on Japanese EFL college learners engaged in a collaborative composition task. It examines the effectiveness of two types of worked examples (those with, and those without, prompts) in relation to the language self-efficacy of the learners within an online collaborative learning environment.

Definition of Terms

The terms and concepts used in the dissertation are defined and operationalized as follows to provide clarity and consistency in their use throughout the arguments and discussions n the dissertation.

Cognitive load. Following the explanation of Paas, Tuovinen, Tabbers, and Van Gerven (2003), it was defined as multidimensional construct representing the load extorted upon tasks or demands required for particular cognitive actions on the leaners' cognitive system. In the dissertation, cognitive load was assumed to be included within an overarching concept, cognitive demand of learning tasks (Niculescu, Cao, & Nijholt, 2009) that is represented by mental effort. Mental effort for on each aspect of three kinds of cognitive loads (intrinsic, extraneous and germane) were measured by three questions (see pp. 54) implemented and tested by Cierniak et al. (2009).

Double-content worked examples. As explained in Renkl et al. (2009), this term was defined as new type of worked examples to guide learners' execution of cognitive actions by two levels of external representations. The double-content worked examples contain two levels of sources of information called domains to be followed, learning

and exemplifying domains. In the dissertation, it included a well-written composition that contained both contents of composition (learning domain) and clear structure of composition (exemplifying domain).

English as a foreign language (EFL). It was defined as a learning environment where English was used as the target language to be learned. English was not used in the daily lives, which was distinguished from ESL (English as a second language) condition where English was used as a medium of communication among the people in the society.

False beginners. As characterized by Mills (2009) and Nakamura (2005), false beginners were defined as the learners retaining low competency in spite of their formal education including their higher education. In the dissertation, the term referred to the college EFL learners who retained low competency after six years of the formal secondary English education.

Language self-efficacy (LSE). It was defined as the confidence in reading and writing of a foreign language. In the dissertation, LSE referred to the degree of confidence in English academic literacy, as well as daily usage. Shaw's (2007) reading and writing self-efficacy scale was chosen to measure LSE. The levels of LSE were thought to reflect the levels of expertise due to past literature suggesting that the level of self-efficacy was closely associated with achievements (Mills, Pajares, & Herron, 2006; 2007; Woodrow, 2011).

Performance. Based on *ESL composition profile* (Jacobs, Zinkgraf, Wormuth, Hartfiel, & Hughey, 1981), the collaborative performance was evaluated by the five criteria of

composition. It referred to the participants' achievement of compositions produced during the online collaboration.

Scaffolding(s). It was defined as a process of assisting the learners to achieve a given task beyond their current capabilities (Dennen, 2004). In the dissertation, both worked examples and prompts acted as scaffolding for the online collaborative composition task.

Stress. It referred to the condition of being distressed by the excessive amount of demands beyond one's coping ability, which was regulated by the level of self-efficacy (Schunk, 1989; 2003). In the dissertation, four types of stressors (Jung et al., 2012) were assumed to cause stress during the online collaborative learning using foreign language.

Stressor(s). It referred to the conditions, agents or other stimulus that could cause stress reactions. In the dissertation, stressors were recognized by four categories based on the study by Jung et al. (2012) that analyzed the stress factors uniquely observed in online collaborative learning environment under EFL context. A term "source of stress" is used interchangeably.

Worked example with prompts. It referred to the worked examples with the addition of prompts, which were short questions and hints that directed learners to their self-explanation of the worked examples. In the dissertation, six self-explanation prompts and three cognitive and meta-cognitive prompts were added to a well-written composition example.

Worked example without prompts. As defined in Renkl (2013), Reisslein, Atkinson,

Seeling, & Reisslein (2006), and Renkl, Stark, Gruber, & Mandl (1998), it referred to an instructional strategy, which showed the problem states and the process of problem solutions, for guiding and assisting leaners during problem solving. Worked example without prompts in the dissertation was the double-content worked example.

Purpose of the Study

The two experiments in the present study aims to investigate the impact of two types of worked examples in an online collaborative learning environment where English is used as a foreign language. The Study 1 attempts to observe whether worked examples with or without prompts can reduce the influence of stressors during online collaboration using foreign language before and after the experiment. Then, Study 2 attempts to reveal how worked examples in online collaborative learning contribute to maintaining low levels of perceived stressors and cognitive loads in relation to levels of language self-efficacy for the participants of low language self-efficacy. Both Study 1 & 2 are interested in how two types of worked examples (worked examples with and without prompts) affect the performance on collaborative composition in an EFL context.

Significance of the Study

The present dissertation is significant in three ways. First, it will clarify the relationships between psychological constructs (language self-efficacy, stressors, and cognitive load) for the learners and online collaborative strategies in an EFL context. The present study hopes to contribute to the development of instructional strategies for

the creation of effective and enjoyable online collaborative environments by investigating how the two types of worked examples could contribute to the reduction of stressors in relation to two levels of self-efficacy.

Second, the findings from the dissertation will serve as the basis for the development of effective instructional strategies to engage learners with low self-efficacy who are assumed to obtain a low level of fundamental knowledge in the learning domain. This will become a tool for educators to facilitate online collaboration for beginners or the learners with limited prerequisite knowledge. The findings will also provide practitioners with insight into how learners from a variety of backgrounds can be taught simultaneously in a liberal learning environment, as well as offer concrete instructional strategies in order to do so. The present study will provide a new perspective that allows us to review and question established theories on stressors and cognitive loads in the context of online collaboration. And it is also aiming at offering pedagogical strategies to college instructors who teach learners with lower language self-efficacy in stressful online learning environments.

Third, the findings from the two experiments will benefit Japanese EFL educators by providing sound instructional strategies for foreign language pedagogy. Japanese EFL educators seeking to implement multi-cultural and multi-lingual online discussions can take advantage of the results. College EFL education in Japan has been the target of criticism for a long time, and accused of being ineffective and unpractical despite a tremendous effort made by practitioners. The present situation, the dissertation will argue, has to do with the incorrect choice of instructional designs and the utilization of inappropriate strategies and EFL pedagogy. In turn, it will posit that the development of learning environment and instructional design for low-level EFL pedagogy is still in its infancy.

The dissertation will discuss and recommend a new direction for EFL

pedagogy in regards to online collaboration in Japan. By applying useful strategies found in this study to their own classrooms, EFL practitioners will be able to arm themselves with appropriate instructional strategies for more empirically sound activities online. Sound instructional strategies will help EFL instructors with minimal knowledge in educational technology when they face the problems of "false beginners", students who have more than six years of formal education, yet still remain beginners in terms of competency level (Mills, 2009; Nakamura, 2005; Thomas, 2006). The dissertation will show how online collaboration combined with new instructional strategies can become a powerful means for effective instruction, especially for those with low-level competency.

CHAPTER 2 LITERATURE REVIEW

This chapter reviews and discusses previous research to highlight key topics that are important for research questions. The chapter first begins with an overview of the fundamental ideas and conceptual frameworks of social constructivism and online collaboration, stressors and cognitive loads and self-efficacy in the EFL context. Then, the discussion moves to instructional strategies that appropriate for learners' level of self-efficacy in relation to stress, stressors and cognitive load theory in the context of online collaboration. In particular, the chapter examines the literature on worked examples with and without prompts when employed as independent variables, as in the two experiments in the present study. The chapter ends by analyzing the issue of under-qualified learners who study English as a foreign language in Japan.

Social Constructivism and Online Collaboration

Online collaboration is one of the key competencies required for learners dealing with the challenges of a rapidly changing environment in the 21st century (Ananiadou & Claro, 2009). The impetus for collaborative learning is rooted in the literature on Social Constructivism. Constructivism assumes that the learning takes place when learners actively internalize learning content and by doing so makes a clear break with the behaviorist paradigm. Criticized for its inefficiency and ineffectiveness, behaviorism sees learning as an individual phenomenon evidenced by a change in behavior as a result of environmental stimuli, while constructivists argue that this definition of learning assumes passive. Gillani, (2003) argues the construction of knowledge that learning takes place when learners actively engage in adding new experience and

interpreting that experience through the lenses of previous knowledge not during passive information reception.

Social constructivists extend the concept of knowledge construction to social interaction, where learners are theorized to build knowledge together through practical learning tasks with peers in a real context (Hsiao, 2005). Social constructivists assume the process of socially constructing knowledge to be significant because it is from these external activities that learners internalize knowledge. The contrast between cognitive constructivism and social constructivism contours, in regards to the site of knowledge construction. The three epistemological views along with instructional positions are compared in Table 2-1 (University of California, Barkley, 2013).

Table 2-1: *The Comparison between Educational Approaches*

	Behaviorists (objectivist)	Cognitive constructivists	Social constructivists
View of knowledge	Repertoire of behavioral responses to environmental stimuli.	Knowledge systems of cognitive structures are actively constructed by learners based on existing structures.	Knowledge is socially constructed.
View of learning	Passive absorption of predefined body of knowledge by learner. Promoted by repetition and positive reinforcement.	Active assimilation and accommodation of new information to existing cognitive structures. Discovery by learners.	Integration of students into knowledge community. Collaborative assimilation and accommodation of new information.
View of motivation	Extrinsic, reward and punishment (positive and negative reinforces).	Intrinsic. Learners set their own goals and motivate themselves to learn.	Intrinsic and extrinsic. Learning goals and motives are determined both by learners and extrinsic rewards provided by the knowledge community.
Implication for teaching	Correct behavioral responses are transmitted by the teacher and absorbed by the students.	The teacher facilitates learning by providing an environment that promotes discovery and assimilation /accommodation.	Collaborative learning is facilitated and guided by the teacher. Group work.

Note. The Comparison between Educational Approaches. Reprinted from *Overview of Learning Theories* by Graduate Student Instructor Teaching & Resource Center. Graduate Student Instructor Teaching & Resource Center, Graduate Division, University of California, Berkeley. Copyright 2013 by GSI Teaching & Resource Center. Reprinted with permission.

In social constructivism, it has been argued that people learn from mediations and scaffolding provided within the Zone of Proximal Development (ZPD) (Vygotsky,

1978), “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (p.86). Vygotsky (1978) has conceptualized two developmental levels to be filled in human developmental stages. One is the level that learners can obtain by their development and efforts. Another is the level that can be achieved with the help of someone or tools that provide them with scaffoldings. ZPD is the zone lined by the two levels. Studying alone does not offer such filling of the distance, thus ZPD should be appropriately bridged or scaffolded (Zang & Peck, 2003) with the help of capable of others for human to construct knowledge. In terms of creating an appropriate learning environment assuring learners’ interactions within ZPD, online collaboration can devise a suitable instructional intervention and liberalized learning context.

For instance, Kurokami, Horita, Yamauchi, Inagaki, and Minowa (2001) argue online collaboration offers a learning condition in which learners become cognitively and affectively engaged. Within online collaboration, learners are encouraged to take authority and responsibility for their own learning process; they become actively committed in the learning process and interactions with peer learners (Kirschner, 2001; Morgan, Whorton, & Gunsalus, 2000). Thus, online collaboration is known to contribute to the development of critical thinking, co-construction of knowledge and meaning, and reflection (Brindley, Walti, & Blaschke, 2009). Online collaboration also encourages deep analysis of the problem, generating more and better ideas as well as quality decisions (Fjermestad, 2004). These and other benefits of actively engaging have gained online collaboration recognition as a great instructional strategy.

Online collaboration is argued to be effective for learners of varied backgrounds. Palloff and Pratt (2005) insist that online collaboration is beneficial for a

wide variety of the population. A community of divergent learners with multiple levels of opinions (Johnson, Johnson, & Smith, 1998) is supported by the "group-centered" environment provided by online learning (Garrison, 2006). Despite these high notes, recent research has revealed that the nature of the online collaboration may in fact pose an obstacle to learning in the form of stressors and cognitive overload.

Stress and Cognitive Loads

A wide variety of studies recognize the negative influence of stress over peoples' courses of action and motivation (Akgun & Ciarrochi, 2003; Lazarus, 1990; Palmer et al., 2003). Depression, poor learning performance and abrupt performance drop are reported as negative consequence of stress by recent studies focused on online collaboration (Jung et al., 2011; Thomée, Eklöf, Gustafsson, Nilsson, & Hagberg, 2007; Zajacova, Lynch, & Espenshade, 2005). These studies suspect stress to be brought on by an excessive amount of pressure from the task or persistent exposure to strong stressors during online collaboration.

Past studies on stress agree that stress is not simply caused by a single pressure originating in the external stimuli. Stress can be generated by multiple sources, including learners' capabilities and experiences. Kudo, Choi, and Jung (2010) has conceptualized the generation of stress in online collaboration as five-factor construct (Technology, Collaboration, Time, Task, and Language) with three kinds of stressors (personal, intrinsic and extraneous intertwined) (see *Figure 2-1*). Individual factors, such as self-control skills, self-efficacy, and social supports are conceptualized to intervene during the generation of stress.

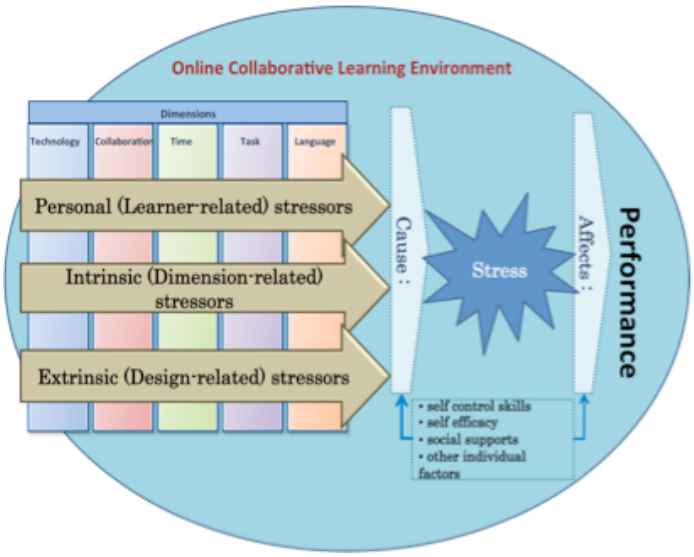

Figure 2-1. The model of stressors in online collaborative learning environment (reprint from Kudo, Choi, & Jung, 2010; p.1002. with permission)

A psychological approach views stress as being "when the perceived pressure exceeds your perceived ability to cope" (Palmer et al., 2003; p.2). According to this view, stress is a result of the imbalance between demands and resources (Lazarus & Folkman, 1984), and therefore is closely related to excessive cognitive demand (Gaillard, 1993; Niculescu, Cao, & Nijholt, 2009). In essence, one feels distress when the perceived demands from a certain task are beyond her capability.

CLT is a body of research focused on 'cognitive demand' or mental effort. CLT research agrees cognitive load is the perceived demand in the cognitive domain during learning. When cognitive demand exceeds the learners' working memory capacity (Miller, 1956), cognitive overload takes place and this hampers learning (Paas, Renkl, & Sweller, 2003; Sweller, 1989, 1999, 2006; Van Merrienboer, Kester, & Paas, 2006). Therefore, it is important to maintain cognitive load within the limit of working memory to assure effective learning (de Jong, 2010; Sweller, 2006; Sweller, van Merriënboer, & Paas, 1998; Kirschner, 2002; Van Merriënboer, & Ayres, 2005).

Niculescu et al. (2009) hypothesized "cognitive demands" to be an overarching construct, thus including both cognitive load and stress. Accordingly, cognitive overload can thus be framed with together with the aforementioned psychological approach to stress (Lazarus & Folkman, 1984). Bandura (1995) further argues stressors that stressors bring stress reactions when one perceives those stressors as being beyond his/her coping ability. The perceived overload can become a source of stress if efficacy for the task is low. On the other hand, if one perceives the stressors as manageable, those stressors will no longer bring about stress reactions.

Therefore, it is assumed that the mitigation of stressors can be realized by keeping demand within one's perceived ability, or self-efficacy, via instructional strategies. Nevertheless, understanding the relation between stress and cognitive load requires more explorative studies to offer empirical evidence.

Cognitive Load Theory (CLT) and Online Collaboration

As discussed in the introduction, management of cognitive load during learning is crucial to ensure the advantages of collaboration. Research on CLT has revolved around instructional designs to minimize cognitive load or when dealing with novel information (Clark, Nguyen, & Sweller, 2006). CLT studies assume that cognitive overload is due to heavy demand resulting from the design of learning tasks (Moreno, 2004). Van Merriënboer and Ayres (2005) champion CLT for yielding efficient learning based on the human cognitive architecture. Clark et al. (2006) describes CLT as "a universal set of learning principles that are proven to result in efficient instructional environments as a consequence of leveraging the human learning processes" (p.7). A body of research agrees learning occurs only when cognitive load is contained within

capacity. Research on cognitive load distinguishes the cognitive demand threefold: intrinsic, extraneous, and germane.

Intrinsic cognitive load is brought about by the intrinsic nature of the task. Paas, Renkl and Sweller (2003) used a term, "element interactivity," to explain the intricate demand generated by the intrinsic nature of the task. They hypothesize that learning tasks may vary in spectrum from low to high in element interactivity. The number of information processing components regulates the level of element interactivity. When element interactivity is high "[e]lements interact and must be processed simultaneously for understanding and learning to occur" (Sweller, 2003; p.218). Element interactivity resides in the complexity of a task, bringing about high cognitive load. The amount of intrinsic cognitive load is determined by the interaction between the learners' expertise level and the number of elements. Studies by Spanjers, van Gog, and van Merriënboer (2012), Pollock, Chandler, and Sweller (2002) and Kudo (2007) argue it is possible to reduce intrinsic cognitive load by decomposing the task and instructional materials, or segmenting the learning process.

Extraneous cognitive load is associated with cognitive processes that, while not directly necessary for learning, manifest due to the complicated design of instructions or learning situations. It is then, by definition, subject to instructional intervention. Learning materials often suffer from unnecessary or overbuilding learning objectives that may produce unnecessary cognitive load for learners. The primary concern of CLT studies is the reduction of cognitive load during learning. Worked examples present a way to reduce extraneous cognitive load.

Germane cognitive load is associated with processes related to learning, schema construction and automation. Recent studies favors experiments aimed at sustaining a high level of germane load (Beckmann, 2010; Paas, & van Gog, 2006). Germane cognitive load should be increased as much as possible without causing

overload. It is essential for instructional designers to manage the total amount of cognitive load to be within learners' cognitive capacity. Recent studies look into instructional strategies to increase germane cognitive load while eliminating extraneous cognitive load. Prompting, used in combination with worked examples, is an option to induce germane cognitive load and will be discussed later (Renkl et al., 2009).

Subjective measurement of difficulty of learning is a reliable and well-tested method for measuring all three cognitive loads mentioned above (Paas, van Merriënboer & Adam, 1994). The method is indirect and had been subject to criticism regarding its reliability as such. Nevertheless, a body of research agrees that people are well capable of comprehending the difficulty of a task and this is considered to be a reliable indicator of cognitive load (Brünken, Plass, & Leutner, 2003). The experiments in the present study employ this method over other options: self-reported invested mental effort, subjective measures (Borg, Bratfisch, & Dornic, 1971; Paas et al., 2003; Paas, Van Merriënboer, & Adam, 1994), self-reported stress level or difficulty of materials (Kalyuga, Chandler, & Sweller, 1999), performance outcome measure (Mayer, 2001), physiological or behavioral measures such as heart rate (Paas & van Merriënboer, 1994) and pupil dilation (Beatty, 1982).

Such studies associate the better learning outcome with lower cognitive load due to sharing of peers in collaboration. Sweller, Ayres, and Kalyuga (2011) exemplify this line of research when they comprehensively discuss the recent findings on collaborative learning and CLT. They summarize findings on the interrelation of task difficulty and shared cognitive load during group learning. They theorize that collaborative learning is advantageous when task difficulty or complexity is exceeded by group capacity whereas, individual learning is more suitable when the task is simple or total cognitive demand is contained within individual limits. They conclude that the effectiveness of collaborative learning depends on the task complexity and the learners'

level of expertise.

Researchers investigating online learning seek a theoretical base in CLT studies (e.g. Van Bruggen, Kirschner, & Jochems, 2002) by focusing on the amount of cognitive load during learning in an attempt to improve collaborative learning. Recent research attempts to overlay the findings of CLT onto collaborative learning to provide reasonable explanations for the observed superiority of group learning. A phenomenon under particular scrutiny is that of expanded working memory shared by participants in collaborative learning (Sweller, Ayres, & Kalyuga, 2011; Kirschner, Paas, & Kirschner, 2009a; 2009b).

Kirschner et al. (2009a) has compared individual and collaborative learning conditions in order to confirm the interaction effect between collaborative learning and cognitive demand. The study has examined the amount of invested mental effort and learning performance on biology tasks. The researchers conclude that individuals involved in group learning invest significantly lower mental effort and perform better on the learning tasks than their counterparts who completed the same tasks alone. They rationalize that collaborators reduce intrinsic cognitive load by sharing and dividing workload, in addition to mental load. They do this by relying on the group as an information processing system, which is called the information reservoir. Thus, collaborative learning makes it possible to moderate the amount of intrinsic cognitive load required for the task. Accordingly, the researchers conclude that collaboration itself has the effect of reducing cognitive load, which they deemed the "collective working memory effect." This is defined as that which allows peers working in collaboration to reduce workload as well as cognitive demand by relying on the group as information processing system and as a result they make the assumption that collaborative peers share the information reservoir and that this contributes to reduction of intrinsic cognitive load.

While it may be assumed that collaboration reduces cognitive load among collaborative peers, by generating collective working memory, this is not without a cost. Additional load, known as "transaction costs," is incurred in the process of interaction among participants in collaboration. Communication among peers could be a load-inducing factor. Kirschner et al. (2009b) illustrates how the elements essential for meaningful collaborative learning instigate transactional costs. In this study it has been understood that in order for knowledge and information to be pooled (shared, discussed and remembered) by participants in collaborative learning, it must be verbalized, justified and reflective upon. These actions generate intrinsic cognitive load.

Kirschner et al. (2011) has recently confirmed a crossover interaction between individual vs. group learning and learning with worked examples vs. problem solving. Comparing the learning outcomes and efficiency on heredity using Dutch high school students. They have found that a load taxing, problem-solving condition is more beneficial for group learning, in contrast to less load-taxing worked examples that are more beneficial for individual learning. The importance of their finding is task complexity is an influencing factor for the choice of individual or collaborative learning. The learners' expertise level as well as task self-efficacy should also subject to choice of learning strategies.

In summary, online collaborative learning appears to be a double-edged sword; on the one hand, meaningful communicative interactions among members can lead to deeper learning and conserve cognitive capacity; while on the other hand, interaction among peers yields substantial working memory exertion, imposing heavy cognitive load. The load imposed intrinsically during collaboration via transactional costs or task difficulties must be balanced by the collective working memory effect, in order for online collaboration to benefit learning. Suitable strategies for online collaboration must address this issue.

Self-efficacy in English as a Foreign Language Context

Hoffman and Schraw (2009) claim that self-efficacy influences the amount of available working memory, thus affecting cognitive competence. Their empirical study on mathematical problem-solving with college undergraduates has revealed that higher self-efficacy correlates to better problem solving efficiency as observed by increased working memory capacity. They insist that self-efficacy affects the size of working memory available for cognitive activities. Language manipulation takes up memory capacity similar to mathematical problem solving, especially when learners are novices or possess limited knowledge. The previous argument is therefore applicable to the use of a foreign language because manipulation of a foreign language requires complex processing made up of cognitive actions that contain interacting elements.

Self-efficacy must be understood if we are to better understand cognitive load. Self-efficacy is defined "as a self-evaluation of one's competence to successfully execute a course of action necessary to reach desired outcomes" (Zajacova et al., 2005; p.678). This definition is grounded on the conceptual understanding that "[p]erceived self-efficacy is concerned with judgments of how well one can execute courses of action required to deal with prospective situations" (Bandura, 1997; p.122).

Past research has found self-efficacy to be influential over human actions and functions as a mediator of behavior. Bandura and Cervone (1986) claim self-efficacy relates to engagement and performance motivation. Detailed analysis was provided by Linnenbrink and Pintrich (2003), highlighting three areas of engagement to be influenced by self-efficacy: (a) behavioral engagement through effort, persistence, and instrumental help-seeking, (b) cognitive engagement through strategy use and metacognition, and (c) motivational engagement through interest, value, and affect. Schunk (2003) adds a mediating function to self-efficacy, arguing that highly

efficacious learners participate more readily, work harder, and persist longer when they encounter difficulties Thus, the level of self-efficacy is assumed to be a determining factor during perception and induction of stress and mental effort in the course of learning.

"Domain-specificity" is an important characteristic of self-efficacy. The concept is that one can have solid self-belief in a particular situation (Pajares, 1996; Pajares & Miller, 1994), but not in others. Mills, Pajares, and Herron, (2006; 2007) examined self-efficacy in foreign language learning and found correlates to higher grades, meta-cognitive strategy use and self-regulation. Woodrow (2011) argues foreign language self-efficacy to be a powerful predictor of writing performance. Mills et al. (2006) find that foreign language self-efficacy relates closely with language proficiency from a correlation study on learners of French as a second language. This investigation allowed them to conclude that highly self-efficacious students perform well in writing tasks in the target language by exerting more effort. Supporting the findings of Mills et al. (2006), Ehrmana, Leaverb, and Oxford's (2003) argue in support for the importance of self-efficacy in English as a second language. Previous research on self-efficacy in language learning uniformly agrees that self-efficacy is an important determinant of achievement, competency and performance in language learning.

The two experiments in the present study choose to evaluate learners' self-efficacy in EFL context via the reading and writing self-efficacy scales developed by Shaw (2007). In her empirical study, Shaw attests to the reliability and accuracy of this scale as a measure for gauging self-efficacy in reading and writing and recommends them for use in research on the relation between ESL learners' self-efficacy and learning outcomes. The scale asks learners to gauge their confidence in skills or tasks on a scale of 0-100. The self-measurement of confidence indicates the level of learners' efficacy for a skill or task in the language. The scale has been originally developed for

high school students in their native tongue; however, the questions are appropriate for low-level EFL learners as well.

Instructional Strategies for Online Collaboration

Instructional strategies refer to the prescriptive means by which instructional conditions are to be arranged, indicating procedures on how to plan those environments in order to achieve instructional goals (Suzuki, 2000). Instructional strategies must eliminate specific problems in order to reach specific educational goals. Without question, in order to develop instructional strategies, there must first be an isolation of problems or deficiencies in the learning conditions instructional strategies must be developed with a specific problem and learning condition in mind.

Recent research has revealed particular problem areas for online learning. For instance, Ingram and Hathorn (2003) argue three principles for effective e-learning: (a) Information presentation, (b) Interactions on the Web, and (c) Connections to other information and experience. Strijbos, Martens, and Jochems (2004) argue for clear guidelines in the design of online collaboration in order to accommodate psychological difficulties such as stress.

Close attention has paid to online collaborative learning has identified four problem areas: (a) the use of technology (Levy, 2009), (b) collaboration itself (McCarthey & McMahon, 1992), (c) time issue (Arnold & Ducate, 2006), and (d) choice of task for collaboration (Mulligan & Garofalo, 2011). Additional research has revealed that using a foreign language can inhibit effective online collaborative learning (Storch, 2005; Lipponen, 1999) when the goal is language learning or when utilized in an international context.

Research interest is growing for the examination of stress found in online collaboration, revealing possible categories of stressors that include: technical aspects and collaboration (Salmon, Allan, & Giles, 2000), technological, organizational and individual aspects (Lawless & Allan, 2004) and lack of trust among the participants (Cohen & Gibson, 2003) These findings has shed light on multiple levels of factors influence performance during online collaboration, especially when used with a foreign language.

Based on research findings on problem areas and potential stressors, a study by Jung et al. (2012) has attempted to identify possible stressors in online collaboration using English as a foreign language in the context of Japanese higher education. The study has collected possible learning problems in online collaboration using a 52 items questionnaire with 226 Japanese EFL learners who have experienced online collaboration. Exploratory factor analysis (EFA) concluded that there were four types of stress factors or stressors: Type I – the lack of self-efficacy, Type II – mismatching with Asian learners' learning styles, Type III – the fear of using online technologies in interactions, and Type IV – the collaborative process itself. Each type of stressor is explained below.

- Type I – A factor for the lack of self-efficacy of the required task covers incompetency or lack of confidence in completion of text-based online discussion using a foreign language, and also efficacy for completing the task at hand. It is natural for online collaborators to be aware of their confidence for presented tasks. The fear of time constraint for task execution is also included, as is linguistic capability when dealing with online asynchronous collaboration. Key items are: Lack of foreign language writing and reading skills, Lack of confidence in a foreign language itself, and Lack of support for

foreign language use during online collaboration.

- Type II – A factor of instructional design for the mismatch of learning styles is related to an ID strategy to select and design the collaborative tasks, as well as facilitating and supporting teaching and learning strategies. The prescribed design or instructional intervention would indicate possible solutions to this factor, especially concerning the structure of tasks. Cultural factors take focus in the discussion of strategy development. Key items for Type II stressor are: Overly difficult tasks, Lack of confidence in posting written format, Lack of confidence in responding quickly, Lack of clear expectation for group tasks and Inappropriate choice of tasks.

- Type III – The fear of technology use is a factor closely related to anxiety or cognitive overload concerning the use of technical tools, courseware and support for both. The use of technology influences the specific learning condition. The fear of technical errors, lack of technical support and difficulties in applying technical skills as well as tools during the learning process applies to this factor. Key items for Type III stressor are: Instructor's lack of timely support, unclear direction for collaborative work, and lack of structure in collaborative work.

- Type IV – Collaborative process itself includes problems in working collaboratively and decision making within learning groups. Collaboration itself naturally involves interactions between peers and instructors, which can be a source of stress. It encompasses managing collaboration, dealing with required learning tasks as well as personal preferences about learning. Key

question items for Type IV stressor are: Lack of technical support, Technical errors caused by online discussion tools, Difficulties in understanding how to use online tools, and Fear of having technical problems.

Another study by Jung et al. (2011) suggests 24 practical instructional strategies designed to deal with the four types of stressors identified by Jung et al. (2012). Table 2-2 shows strategies that mitigate the effect of stressors in and out of online collaboration. The matrix addresses how these strategies can mitigate influence of stressors by attending to the four stress factors as well as the stage of collaboration. These strategies have yet to be validated empirically and therefore are subject to criticism for being untested. As such, detailed and empirical examinations are required to confirm the effectiveness of these instructional strategies. Without such empirical evidence, no claims can be made regarding the use of these strategies in realizing effective online collaboration using a foreign language.

Table 2-2

Strategies to Reduce Stress of Asian Learners in a High Context Culture in Online Collaboration Using a Foreign Language

Four principles Timeline	1) Promoting self-efficacy	2) Matching design strategies with learning style	3) Reducing technology- related fear	4) Facilitating collaborative process
Prior to online collaborative session	- Analyze students' prior learning experiences and language competencies - Plan support for learners with different experiences and language competencies	- Choose appropriate tasks to learners' needs and learning objectives - Design collaboration structured with specific objectives & outcome statements - Estimate precise & realistic time table for collaboration - Design reward system related to the desired behaviors	- Provide short online technology hands-on training - Offer personal technology support to learners with low technology skills	- Take account of collective cultural expectation not to lose the face - Set clear rules, regarding types of collaborative activities and assessments, for collaboration that are comprehended by all the learners - Create heterogeneous groups - Assign a facilitator and clearly present individual roles during online collaboration
During & Initial stage of collaboration	- Provide frequent encouragement with positive feedback to individual learners - Offer frequent & timely group support during collaborative group work	- Provide instructor-led activities offering clear direction & expectation for group task aims, protocols & procedures at the initial stage - Introduce more learner-directed collaboration after the initial, instructor-led activities	- Set up an 'just-in-time' online support system (e.g. FAQ) using simple language for both technical problems & problems related to collaboration process - Offer on-going technology support when needed	- Assign clear role(s) to each group member for a collaborative task - Facilitate group activities or encourage active facilitation of the assigned facilitators - Promote social interactions by providing spaces for social interactions, encouraging profile exchanges, or online *getting acquainted* or *ice breaking* activities
Evaluation		- Promote self-reflection of own collaboration through journal writing		- Combine individual & group assessments - Set clear evaluation criteria

Note. Adopted from Jung, I.S., Kudo, M., & Choi, S.K. (2011). Instructional design strategies for stress-reduced online collaboration in Asia's high context culture. *Proceedings of the 25th annual AAOU conference,* p. 55.

Matching Instructional Strategies with Learners' Level of Self-efficacy

The argument makes it clear that stress found in online collaborative learning can be treated by category with specific instructional strategies to mitigate them. The study by Jung et al. (2011) illustrates the need for close attention to the Type I stressor discussed in the previous section, namely: lack of confidence and efficacy in fundamental task completion for online collaboration. They insist instructional strategies are needed to scaffold learners' lack of language related self-efficacy. Their study has revealed that online collaborators are overwhelmed by cognitive demand from the collaborative tasks, resulting in considerable loss of effectiveness during collaboration using EFL. Jung et al. has recommended scaffolding both the collaborative task and processes to prevent an over demand from cognitive tasks, especially when learners have limited background knowledge or exhibit weaker language self-efficacy.

Bannert (2002), Jung et al. (in press), and Kudo (2013) recommend the use of worked examples (Reisslein, Atkinson, Seeling, & Reisslein, 2006; Renkl, Stark, Gruber, & Mandl, 1998) supported and developed by CLT literature as a means of instructional strategy to scaffold self-efficacy during online collaboration. Worked examples are known to assist learners in avoiding means-end problem solving procedure, and are shown to contribute to reduction of cognitive load during the learning process (Atkinson et al., 2000; Renkl, 1997, 2002, 2005). A recent CLT study posits that worked examples can supplement lacking knowledge by providing problem solution (Hübner et al., 2010). The provision of worked examples is gaining ground as an instructional strategy for scaffolding learners' poor linguistic competency as well as cognitive skills. This will be discussed later.

Worked Examples as the Instructional Strategy

Van Merriënboer and Ayres (2005) summarize six CLT-driven effects based on a meta-analysis of CLT research: (1) the goal-free effect, (2) the worked example effect, (3) the completion problem effect, (4) the split attention effect, (5) the modality effect, and (6) the redundancy effect. These effects are benefits from the design of problem formats. The formats and the effects they bring about offer a unique perspective on cognitive load for the instructional strategist. This dissertation focuses exclusively on the worked example effect.

Sweller, Van Merrienboer, and Paas (1998) provide a classic instance of worked examples. The worked example consists of "a problem formulation, solution steps, and the final solution itself" (Renkl, 2002; p.529). Because of the format of worked examples and worked examples help leaners find a step by step procedures for problem solving through the observation of examples (Sweller, 1999), they are known to be effective for novices or learners lacking prior knowledge (van Gog & Rummel, 2010; Van Merriënboer & Sweller, 2005). Figure 2-2 describes the classic worked example by Sweller, Van Merrienboer, and Paas (1998).

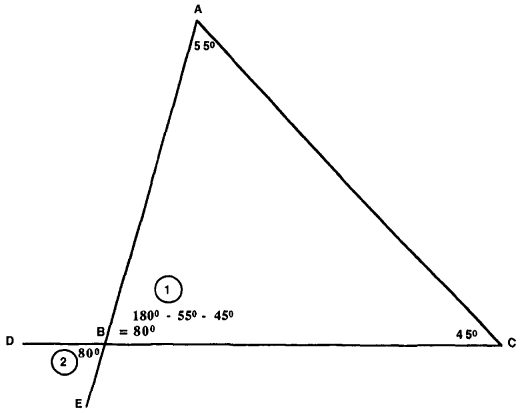

Figure 2-2. A worked example. Adapted from Sweller, J., Van Merrienboer, J. & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), p. 279.

In addition to increasing efficiency, worked examples are known to be effective for problem solving. Reisslein et al. (2006) empirically has tested the worked example effect by demonstrating the problem solving procedure and concluded the use of worked examples contributes to the reduction of cognitive load. They compare the benefits of three different computer-based learning environments, in which worked examples is followed by problems; problems are followed by worked examples; and a fading problem format where worked examples faded away along with the progression. They find the example-problem format advantageous for lower achievers, while problem-example format effective for higher achievers. This finding has become a base for the hypothesis of the worked example effect and the expertise reversal effect (Kalyuga, 2007; Kalyuga et al., 2003; Tobias, 1987) that same format of instruction can be effective for the lower level learners, but redundant and hampering for higher level learners.

Sweller (2006; 2009) reports that learners using worked examples reduce cognitive load thorough observation. By observing, a learner is able to borrow a part of example to complete her own task, consuming far less working memory (Sweller & Sweller, 2006). CLT research has theorized that learners reduce mental effort through observation, examination, borrowing and reorganization of the examples, thus eliminating the risk of randomly applying problem solving through trial and error or means-ends analysis (Atkinson et al., 2000; Chandler & Sweller, 1991; Renkl et al., 1998; Sweller, 2006; Sweller & Levine, 1982; van Gog & Rummel, 2010).

The Use of Worked Examples under Online Collaboration

The study by Retnowati, Ayres, and Sweller (2010) implement worked examples in

online collaboration and find there is no particular difference between individual and collaborative learning in the effectiveness of worked examples for managing cognitive load. They have compared learning conditions using worked examples and problem-solving in collaborative and individual learning context. Comparing 2 (problem-solving vs. worked example) x 2 (individual vs. group study) with Indonesian 7th graders in a mathematical domain, the study have found that both individual and group study conditions benefit more from worked examples than from problem-solving alone. This is credited to reduction of extraneous cognitive load by the use of worked examples. They have not confirmed the advantage of group learning over individual learning that Kirschner et al. (2009b) has found.

Krause, Stark, and Mandl (2009) has verified the benefits of worked examples in a similar investigation of group vs. individual learning. They used an example-based e-learning system for statistics and compared outcomes between individual and cooperative learning. The study finds that cooperative learning itself does not lead to greater learning outcomes. Cooperative learning does, however, promote better performance on learning tasks. This has been measured by the results of six problem-solving tasks on the e-learning system, which has shown to be superior after the treatment. The worked example has thus been used as an instructional strategy and functioned well, combined with or without feedback intervention.

Double-content Worked Examples

Traditionally, the application of worked examples has been effective in algorithmic fields of study such as algebra, (Sweller & Cooper, 1985), geometry (Paas & Merriënboer, 1994; Tarmizi & Sweller, 1988), physics (Ward & Sweller, 1990),

chemistry (Crippen & Earl, 2004, 2007), and computer programming (Trafton & Reiser, 1993). However, recent effort has opened up the possibility of application to non-algorithmic fields of study such as design education (Rourke & Sweller, 2009), argumentation skills acquisition (Schworm & Renkl, 2007), promoting collaborative capabilities (Rummel & Spada, 2005), cognitive skill acquisition (Atkinson et al., 2000) and strategies for learning journal composition (Hübner et al., 2010).

There have been attempts to develop new type of worked example for learning heuristics beyond algorithmic fields of learning. These new worked examples are assumed to be versatile for non-algorithmic fields; the past research on worked examples demonstrated the effect seen in algorithmic fields of studies expanding to include ill-structured instructional areas (Renkl, 2005). The use of worked examples is "equally effective irrespective whether the problems are well-defined or ill-defined" (Rourke & Sweller, 2009; p.187).

Double-content worked examples are distinguished from the classic worked example; classic worked examples have a single content, while double-content worked examples provide two levels of example within the format (Ayres & Paas, 2009; Renkl et al., 2009; Stark, Kopp, & Fischer, 2011). Renkl et al. (2009) describes the new format of worked examples as being effective for non-algorithmic domains. They explain that double-content worked examples have learners process two levels of domain: (a) learning domain (the structural patterns to learn) and (b) exemplifying domain (the topic of composition for observation), which correspond to two characteristics pointed out in the previous section, modeling by socio-cultural learning theory. In a way, double-content worked examples are a modeling without well-trained procedural instructions.

Rourke and Sweller (2009) attempt to utilize worked examples for the benefit of advanced learning performance in a non-algorithmic domain. They implement

worked examples for identifying chair designers' signature styles. The learners have charged with learning to differentiate styles of chair design by studying worked examples. The worked examples used in the study include the problem and key descriptions indicating distinctive features of each chair. The study has compared the learning performance of a worked example group with that of a problem-solving group. The researchers conclude that the double-content worked example group has outperformed the problem-solving group in tasks for novice learners with a moderate level of visual literacy skill.

Renkl et al. (2009) also explains that a well-written composition example, a kind of double-content worked example, must consist of two levels of contents for the leaner if she is to benefit from it. Learners are to observe and model two levels of contents, namely: such as "the structure of the composition (learning domain)" and "the meanings of the writing (exemplifying domain)." as a topic so they can benefit out of the example. By seeing both logical structure and the semantics in a in the written format as of an example, the learners can benefit from two domains of content. This is less load-taxing than being explained and reading an explanation on how to compose a good writing. Renkl et al. (2009) warn, however, that neither learning nor borrowing may take place if learners have difficulty in comprehending the contents of the exemplifying domain.

Among developmental studies of worked examples for heuristic learning, Hilbert and Renkl (2009) implement the double-content (sales contracts and amortization) example for heuristic learning to acquire a learning strategy called concept mapping. They investigate the effects of the new type of worked example in comparison with practicing. The first experiment has been a comparison between a worked example group and a practicing group. They find no advantage for the worked example group. The second experiment, which adds another group with self-explaining

prompts added to the worked example, has shown better learning outcomes. The addition of prompts to the worked example has been confirmed effective for the heuristic learning domain, with a caveat: a sufficient amount of self-explanation is required because an example alone might be too difficult to understand.

Most recently, Kyun, Kalyuga, and Sweller (2013) experiment with the effectiveness of double-content worked examples for English essay writing with learners in Korea. The participants have been shown three well-written double-content worked examples as worked examples for English literature. The researchers have explained that the leaners used genesis and the borrowing principle (Sweller, 2006) and expended less cognitive effort than the group practice without worked examples. They find out that the effectiveness of worked examples is more pronounced when they are used for the learners of lower English proficiency.

In spite of this promising research, double-content worked examples with or without prompts have yet to be tested in online collaboration. The next section reviews attempts to expand the possibility of worked examples by adding prompts appropriate for the induction of proper learning strategies or implicit instructions. Prompts are considered suitable for online collaboration because here learners are more autonomous in task completion.

Worked examples without prompts – the importance of self-explanation. Study 1 and 2 in the present study has utilized double-content worked examples. The worked example without prompts group used a double-content worked example. The worked example with prompts group has utilized the same double-content worked example with prompts included.

As argued in the previous section, recent research (e.g., Hilbert & Renkl, 2009; Rourke & Sweller, 2009) is occupied with improving the format of worked

examples. However, the utilization of this new type of worked example does not in itself assure greater learning outcomes. Koedinger and Aleven (2007) posits the self-explanation process to be crucial when learning with worked examples. Learners are apt to suffer from the illusion of understanding (cf. Chi et al., 1989; Bannert, 2002). A study by Chi et al. (1989) states that "poor learners of physics skim worked-examples in textbooks and make shallow analogies when solving homework problems whereas good learners try to explain to themselves the reasoning from one step to the next and then make deeper analogies during problem-solving practice" (p.257). For worked examples to be highly effective, the instructor must provide learners with the opportunity to self-explain or elaborate on the example itself.

Renkl (1997), experimenting on Dutch secondary education students, has compared qualitative differences in self-explanations and has discovered that superficial or passive self-explanations are not effective. The work concludes the quality of self-explanations is a substantial determining factor for learning with worked examples, thus deeper self-explanation should be promoted. The use of prompts in order to successfully induce meaningful self-explanations is gaining particular attention from the research community (Atkinson, Renkl, & Merrill, 2003; Renkl 2005), especially the combination of double-content worked examples and prompts (Van Gog & Rummel, 2010).

Worked examples with prompts. The research has proven adding questions and short directions onto worked examples induces more effective self-explanation, leading to better learning outcomes and learning strategies (Reigeluth & Stein, 1983), in addition to fostering transfer ability (Renkl, 2005). Prompting is a technique to induce self- explanations through short questions or elicitations (Schworm & Renkl, 2007). A number of laboratory studies integrate prompts with worked examples, yielding

qualitatively superior learning results (Atkinson et al., 2003; Chi, de Leeuw, Chiu, & LaVancher, 1994; Renkl et al., 1998). The prompts are thought to induce deep elaboration of the example using metacognitive strategies and encourage carefully reading of the compositional elements and increase awareness of suitable expressions.

Schworm and Renkl (2007) have compared worked examples with or without prompts, highlighting the quality of self-explanation. They examined 71 student teachers and have found an increase in argumentation skills when prompted. Similarly, Hübner et al. (2010) compare worked examples with or without prompts with German high school students. These novice students have engaged in learning post-class journal writing. The worked examples used in their experiment is a double-content worked example, which the researchers describe as "an example of a well-written learning journal, showing the critical features that make a learning journal well-written (e.g., well-organised, including many examples, high amount of meta-cognitive statements, etc.)" (p.21). The provision of a well-written example for composition tasks has been confirmed effective for supporting deficiency in writing competency (Renkl et al., 2009). Furthermore, the worked examples with prompts group have outperformed all other groups in a transfer session conducted seven days later; the outperformance have been observed in the experiment on the transfer problems to another context. The study concludes that the worked example with prompts group has generated better self-explanation and the combination of worked example and prompts have served to scaffold the leaners of lack of previous knowledge. Prompts have fostered learning strategies such as self-explanation, which consequently has lead to effective learning outcomes (Chi et al., 1989).

The findings above confirm that prompts enable effective and efficient learning when combined with the new type of worked example (Renkl et al., 2009; van Gog & Rummel, 2010). Additional studies (e.g., Hilbert et al., 2008; Renkl et al., 2009)

however, raise a concern; the self-explanation induced by prompting may overload novice learners when they learn complex content. Self-explanation generates germane cognitive load and this proves detrimental leads to cognitive overload (Kalyuga, 2008; Renkl et al., 2009; Sweller, 2006). It is still under debate whether prompts lend an extra hand or a concrete scaffold to novice learners or not, especially for lower achieving EFL learners engaged in collaborative learning. This is especially relevant to the online implementation of collaborative learning, where learners often suffer from cognitive overload due to lack of skills and prerequisite knowledge.

Prompts have benefits beyond the induction of learning strategy. Davis (2003) claims that prompts can act like an advisor "directing" the learner to self-explain worked examples. In her study with high school students possessing limited pre-existing knowledge, prompts became an extra pair of hands for learners' with limited competency. The metacognitive prompts used in the study may act as a "more able other," and are closely related to directed prompts (Davis, 2003, p.15). This explanation is in line with Renkl (1997), who explicates that co-leaners' questions during collaboration motivate and improve learning for less motivated learners; he posits that these questions are equivalent to prompts. Such hints go beyond inducing strategy, to support motivation for learners with low self-efficacy.

The College EFL Leaners in Japan Today

This study is strongly concerned with the present state of college EFL education in Japan. Two points should be made in order to grasp the characteristics of EFL leaners in Japan today.

Non-traditional college students. Guldberg and Pilkington (2006) argue for the existence of an issue with non-traditional students in the university. Although there is currently no agreed-upon definition for non-traditional college students, the western idea of non-traditional college students includes the elderly, single parents, part-timers and those without high school diploma. Non-traditional college students in Japan are slightly different from those in the western context.

What Japan is experiencing is very similar to what Cole (2009) identified in England, namely: a change in learner quality and increasing numbers of incompetent learners at the level of higher education. These under-qualified students are incompetent in their studies and often demonstrate low self-efficacy.

The growing numbers of under-qualified learners in Japanese higher education is synchronized with the fertility decline in Japan. Numerous Japanese colleges are under-enrolled (Shepherd, 2008). Colleges compete for enrollment and scores of students are admitted who would not have made the cut in previous years. This only increases the number of Japanese non-traditional students and the problem is growing every year. This is especially true for colleges located far from the capital, Tokyo. These colleges are experiencing extreme difficulties with students recruitment.

The issue of under-qualified students in Japanese higher education is widely discussed. The focus of attention is on how under-qualified college learners tend to lack sufficient motivation, engagement and persistence. The researcher also finds more and more students lacking sufficient motivation, absent from class, neglecting class requirements, dropping credits or even dropping out of college altogether in recent years. Non-traditional college learners do not have proper motivation to learn at college.

EFL education in Japan. EFL instructors in Japan have long discussed the low competency of learners, (especially engineering majors), their ineffective

performance (Kameyama, 2007) and lack of motivation (Johnson, 2010).

Helgesen (1991) confirms that numerous college students cannot even maintain a simple conversation in English. Kanatani (2004), Takefuta and Suiko (2005), and Yoshida (2004) frame the issue as two pedagogical dilemmas for EFL education in Japan: (1) the extensive use of traditional instructional strategy, the Grammar-Translation method (GT method); and (2) inadequate amount of training to be competent in English (Asher, 1972; Morley, 1990). EFL learners need more contextualized and learner-centered instructional design, making use of output, interactions, reflection, and evaluation among learners in the target language. Despite these developments in the discussion on Japanese EFL, the problems mentioned above still manifest in pedagogical practice across the country.

Another factor contributing to very low competency in EFL education in Japan is the "demotivation" of English language pointed out by recent studies (Arai, 2004; Dörnyei, 2001) This is a result of negative experiences from old-fashioned language pedagogy or unsuccessful learning experiences in English education. It has been widely discussed that general dislike of English begins to manifest at the level of secondary education. Kudo et al. (2003) attests that there are increasing numbers of high school students who feel nervous in learning or have never experienced a sense of achievement in their study of English. Similar observations are reported at the college level. Studies confirm that past learning experience contributes negatively to the dislike of English or low motivation for Japanese college EFL learners. Students majoring in technology display a particular dislike of English (Kameyama, 2007). Johnson (2010) as well as Kimura, Nakata, and Okumura (2001) report that technology majors suffer from low achievement in English because of a dislike for the language and low competency. This trend began with the introduction of remedial training which symbolizes the educational reform efforts instigated in recent years.

The Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) respond to strong demands from society and submitted instructional guidelines in 1999 to increase the practical training of English, emphasizing "oral communication skills". The guidelines advocate balancing out the grammar centered learning with more practical oral skills training. The gesture has made a great impact on EFL education in Japan, signaling a shift to more practical pedagogies and exercises based on a learner-centered approach. This has led to gains in competence for a section of the student population. It has been ineffective, however, for the rest of learners; despite the huge effort, overall communicative competence remains low. Scholars have begun to call these students "false beginners" because they have more than 6 years of formal education, yet still remain beginners in terms of competence (Mills, 2009; Nakamura 2005) with low interests over English as a school subject and low motivation from EFL condition.

Research Questions

As seen above, previous studies pointed out the necessity of sound instructional strategies for maintaining a proper level of stress by understanding the nature of stressors in online collaboration, especially in the EFL context. While these studies identified major stressors during online collaboration, effective instructional strategies for improving the online collaborative learning experience have yet to surface. While worked examples are one effective instructional strategy that contributes to the minimization of extraneous cognitive load, they have not been examined as a means to mitigate the influence of stressors. In addition, while there are several claims that these two constructs - cognitive load and stress - are closely related, empirical evidences is

scarce, especially in the online collaborative learning environment where EFL is used as a communication tool.

Thus, this study examines worked examples as an instructional strategy to mitigate the influence of both stress and cognitive loads in EFL-based online collaborative learning. It focuses on two types of worked examples (worked examples with prompts and worked examples without prompts) and investigates the effects of these two types of examples on three variables: self-evaluated stressors, perceived cognitive loads, and the quality of collaborative composition, while considering the levels of learners' language self-efficacy. Specific research questions are as follows:

1. What type of worked examples would be effective in reducing stressors of EFL learners with two different levels of language self-efficacy in an online collaborative learning context?

    a. Does the provision of worked examples with prompts mitigate Type I stressor (a stress factor related to the lack of self-efficacy) of EFL learners with higher and lower levels of language self-efficacy?
    b. Does the provision of worked examples without prompts mitigate Type I stressor of EFL learners with higher and lower levels of language self-efficacy?
    c. Which type of worked examples - with or without prompts - is more effective in mitigating self-evaluation of Type I stressor of EFL learners with higher and lower levels of language self-efficacy?
    d. How do the other three types of stressors (Type II – mismatching with Asian learners' learning styles, Type III – the fear of using online technologies in interactions, and Type IV – the collaborative process itself) change by implementing two types of worked examples in online collaboration?

2. What type of worked examples would be effective in reducing cognitive loads of EFL learners with two different levels of language self-efficacy in an online collaborative learning context?

   a. Does the provision of worked examples with prompts mitigate cognitive loads (intrinsic, extraneous, and germane) of EFL learners with higher and lower levels of language self-efficacy?
   b. Does the provision of worked examples without prompts mitigate cognitive loads of EFL learners with higher and lower levels of language self-efficacy?
   c. Which type of worked examples - with or without prompts - is more effective in mitigating self-evaluation of cognitive loads of EFL learners with higher and lower levels of language self-efficacy?

3. What type of worked examples would be effective in contributing better collaborative performance of EFL learners in an online collaborative learning context?

CHAPTER 3 METHODOLOGY

This chapter describes the methodology employed in the two experiments in the dissertation. Two experiments were conducted in the year 2012 with samples of different age groups in the same department of the same institute to ensure the findings of both studies.

Study 1 was a quasi-experiment conducted in an institute of technology in Japan, using two intact groups. The experiment was carried out for six weeks with 60 junior (3rd year) students to observe the effects of two instructional strategies, worked examples with and without prompts over four types of stressors (Type I – the lack of language self-efficacy, Type II – mismatching with Asian learners' learning styles, Type III – the fear of using online technologies in interactions, and Type IV – the collaborative process itself), three kinds of cognitive loads (intrinsic, extraneous, and germane), and performance on collaborative compositions. The levels of language self-efficacy were considered to be an intervening variable that mediates the influence of treatments upon stressors, cognitive loads and collaborative composition performance for participants in online collaboration in EFL context.

Study 2 added a control group in addition to the two treatment groups. Employing a control group allows more precise comparison between the effects of instructional strategies. Study 2 also employed a quasi-experimental method using four intact classes in the same institute as in Study 1. Study 2 used lower-grade college learners (80 freshmen and sophomore students – 1st and 2nd year) compared with the participants in Study 1. These students had less experience in college English learning and online collaboration. Study 2 carried out the same six-week experiment to investigate the effects of worked examples with and without prompts over four types of stressors, three kinds of cognitive loads, and performance on collaborative compositions.

As in Study 1, levels of language self-efficacy were examined as an intervening variable for stressors and cognitive loads.

Study Focus

Study 1 focused on how four stressors, cognitive loads and collaborative composition performance as discussed in Chapter 2 would be affected by the use of two types of worked example as instructional strategy before and after the treatment during online collaborative composition performance. It attempted to observe the change of stressors before and after the treatment. Study 1 was an exploratory attempt to observe the anticipated effects of the treatments. Consequently, it fulfilled the role of manipulation check.

Study 2 was an attempt to reflect the findings and recommendations proposed by Study 1 and collect further empirical evidence for a firm understanding of the effects of worked examples with and without prompts in the online collaborative learning environment. The results from Study 1 indicated that lower language self-efficacy learners benefitted more from the additional prompts used as instructional strategy. Study 2 was administered as a similar experiment with participants assumed to be of lower foreign language competency and less experienced with collaborative tasks that college education nowadays often utilize.

Although the frame of experiment was identical, some adjustments were made. Study 2 used four intact classes, and increased the number of participants. The four intact classes were divided into three groups by random grouping as per the three conditions. While Study 1 was a mere comparison between the treatments, Study 2 employed a control group that provided no instructional strategy during the

experimental period. Also, Study 2 reduced the number of participants of collaborative team from triads to dyads, reflecting the recommendation from Study 1. Study 1 suffered from problems of loafing and dropouts due to the grouping of participants.

Context of the Study

The context of both studies was a regional Japanese town where the learners have no imperative needs for English in daily lives. The condition is called EFL and is distinguished from the condition of English as a Second Language (ESL), which takes place in English speaking country. Without a context for practical use, English is nothing but a required subject to pass at school.

The specific site of the studies was a private technical institute located in Sapporo, Japan. The institute has over 2500 students majoring in technology. The colleges in the local area often gather "under-qualified" students because of under-enrollment as discussed in Chapter 2. Working at a regional institute of technology, the researcher has recognized the current problem of increased numbers of non-traditional students enrolling, in addition to a problem of low motivation and competency. These problems are getting worse year by year.

Both studies were carried out in a department called Media Communication Design, in which students are trained to be web designers, game creators or system engineers. As discussed in Chapter 2, technology majors are known for not being good at English, or even disliking English. There are students in the class with low English competency who are not able to form a short simple sentence in English.

Participants

Participants in Study 1. Participants of Study 1 were juniors. The gender balance was 82% male and 18% female. The average age of the 60 undergraduate students was 20.5. They enrolled in the class, Internet English I as an elective. The course objective is to foster Internet related English competences. The participants joined in the collaborative composition task to fulfill the course requirement of a final project. The task comprised 30% of the course evaluation and was judged on quality of collaborative composition, participation as well as effort. The instructor used a textbook, *First Steps in Academic Writing* (Hogue, 1996), to teach structured writing.

Majoring in a department of ICT related area, the participants were well informed of the current Internet services, tools and related operations. All the participants owned personal computers and used them in and out of the classroom. The school provided all students with off-campus Internet connection via a virtual private network (VPN), allowing access to the intra-network of the college. The advantage of using an online collaborative learning environment is that the students were able to work on the collaboration at their own convenience and apply recent developments of mobile as well as ICT technologies.

During Study 1, the students formed 20 triads via the random group assignment function of the Moodle system. Two treatment groups, worked example with prompts or worked example without prompts, were formed with the intact classes (hereafter class A and B) of Internet English I. Class A had a total of 11 triads used worked example without prompts. Class B had 9 triads that used worked example with prompts as instructional strategy. At the end of the experiment, 8 triads did not complete all procedures. This was due to the loss of communication during the experimental period or students dropping out of the class because of low learning

motivation. As a result, a total of 12 triads remained after the experiment. This drop out rate is not unusual for a course in a low-level college, which gathers less motivated leaners.

Participants in Study 2. The academic backgrounds and characteristics of the participants of Study 2 were similar to Study 1. The average age of the 80 Japanese undergraduate student participants was 19.1. The gender balance was 78% male and 22% female that was almost identical to the Study 1. One major difference was the years of enrollment. Study 2 employed as participants freshmen and sophomores at the same college as Study 1. In Study 2, the participants enrolled in the English classes a required course at the general education level. The course objectives were to foster basic English reading and writing competences. All participants had low knowledge of English composition structure prior to the experiment. They joined in the collaborative composition task to fulfill the course requirement of a final project. The task comprised 30% of the course evaluation and was judged on quality of composition, participation as well as effort. The same textbook, *First Steps in Academic Writing* (Hogue, 1996), was used to study writing procedures.

The participants were the students from 4 intact classes: 18 from class A, 16 from class B, 18 from class C, and 14 from class D. Within the intact classes, the participants were assigned to either one of the three treatment groups: control, worked example with prompts and worked example without prompts. In Study 2, the collaborative teams formed dyads, reflecting the recommendations from Study 1. The participants were grouped with their partner by the random grouping function of the Moodle system and randomly assigned into the treatment groups using the random table (JUCE). This allows us to assume that the groups were equal in makeup. At the end of experiment, 7 dyads did not complete the collaborative composition or finish the tasks.

As a result, 33 dyads took part in the experiment.

Collaborative Task

To assure practical writing training and interaction among the participants, three guiding principles were established: (1) the task must be too difficult to complete alone, but made possible by collaborative work that assists task completion through division of labor; (2) the task must be relevant to the participants' lives in order to maintain a common goal for both learners and the instructor and (3) enough time should be allotted for task completion so that low achieving learners can complete the collaborative composition task.

In both Study 1 and 2, promoting collaborative learning among the participants through suitable selection and organization of the collaborative task was considered important, especially for participants of low English competency. Jung et al. (2011) analyzed stress factors for online collaboration and found that suitable tasks for EFL learners in an online collaborative learning context should be just beyond the effort an individual is capable of. A task simple enough for the individual does not require collaborative effort and collaboration itself can be a stress generator (Kirschner et al., 2009a).

For both studies, collaborative composition task was chosen as a means for practical training for false beginners. It was expected that even low-level EFL learners can contribute to the task by sharing cognitive demand with peers. Online collaboration provides learners with opportunities for using English as a means of communication by utilizing a practical context. Also, learners are able to practice generating output with other peers who may co-edit incomplete compositions together.

The collaborative task in both studies was to compose a letter addressed to high school students. In the letter, each group had to recommend Sapporo as a location for college using casual English expressions so that high school students can understand without difficulty. Compositing a letter in English for high school students is a topic to which college students can easily relate.

A double-content worked example, with or without prompts, was provided during the collaborative task, according to the treatments as instructional strategy. The example was written by a native speaker of English who engaged in English education in Japan. Two professionals reviewed the example in terms of readability and difficulty of expressions. During collaborative composition task, communication related to task completion was allowed in Japanese, but the composition was to be written in English.

Instruments

Instructional strategy. Both Study 1 and Study 2 used two variations of an instructional strategy: worked examples with prompts and worked examples without prompts.

Worked example (WE) without prompts. Participants in the WE without prompts group were provided with a double-content worked example only. The double-content worked example in both Study 1 and 2 was a well-written composition example, containing a well-structured logical organization, good illustration and support of ideas, deep analysis of problem and appropriate choice of words as discussed in Chapter 2. The important structure of composition, such as transitions, topic sentence, supporting sentences, etc., were highlighted visually by the different colored fonts and types of fonts. The participants were provided with a two-page long model letter and

were able to refer to this worked example freely during the experimental session. The content of the letter was a recommendation for high school students to go to colleges in Nagano instead of Sapporo. The WE without prompts is attached in the Appendix C.

WE with prompts. Using an identical double-content worked example, nine self-explanation prompts were inserted at the end of the worked example. Six questions were used to induce reflection on their own compositions.

1. "What is the title of the composition?"
2. "Does this writing have topic sentences? What kind of information was provided?"
3. "How many supporting sentences does this composition have? What do they explain? Do you find transitions?"
4. "How many 'reasons' were there? Which sentence shows the reasons?"
5. "How many examples does each reason have?"
6. "What are the concluding sentences for the letter and each paragraph?"

Three prompts were added to induce applicable cognitive and metacognitive strategies for worked example following Berthold et al. (2007). The WE with prompts is attached in the Appendix D.

1. "How can I best organize the structure of the learning content?" (cognitive – organizational stimulation);
2. "Which examples can I think of that illustrate, confirm or conflict with the learning contents?" (cognitive – elaboration);

3. “Which main points have I already persuade high school students well?” (positive) and “Which main points haven’t I persuade high school students yet?” (negative) (meta-cognitive – monitoring and self-regulation).

Language self-efficacy (LSE). To measure participants’ LSE, Shaw’s (2007) reading and writing self-efficacy scale was chosen for both Study 1 and 2. Shell, Colvin, and Bruning (1995) has developed the scale, and it is widely used. Shaw attests to its validity and reliability as a measurement of LSE. It is appropriate for measuring EFL learners’ LSE because the scale is divided into two segments, reading and writing parts. Both of these parts are sub-divided into basic skills and day-to-day simple foreign language-related tasks. Each segment of the scale contains 18 items addressing task-efficacy for reading, 11 items addressing reading skill-efficacy, 16 items addressing task-efficacy for writing, and nine items addressing writing skill-efficacy. The participants rate their confidence for each item on a scale of zero to 100. It was translated into Japanese and checked by three professionals for accuracy of the translation.

Upon analyzing levels of LSE in Study 1, participants (n=36) were divided into two groups. The participants were first divided into the three groups by the method of a good-poor analysis centered by the .20 of the one standard deviation (*SD*) from the mean score. The middle group (n=4) was excluded and the mean was taken from the high and low groups. The lower range of the high LSE group (n=16) was determined by adding .20 of *SD* onto the mean score. The higher range of the low group (n=15) was determined by subtracting .20 of *SD* from the mean score.

Participants (n=66) in Study 2 were divided into two groups. The participants were first divided into the three groups by the method of a good-poor analysis centered

by the .20 of the one *SD* from the mean score. The middle group (*n*=12) was excluded and the mean was taken from the high and low groups. The lower range of the high LSE group (*n*=27) was determined by adding .20 of *SD* onto the mean score. The higher range of the low group (*n*=27) was determined by subtracting .20 of *SD* from the mean score.

Perceived stressors. Learners' perceived stressors were measured by a stressor scale developed by Jung et al. (2012). Even though this scale was recently developed and does not have accumulated reliability evidences, it identifies four stress factors particularly affecting Japanese students' online collaboration based on empirical data. The scale is divided into four types of stress factors or stressors. In total, there are 32 items in the scale: 12 items for confidence and language – Type I stressor, seven items for design of instructional environment – Type II stressor, six items for the use of technology – Type III stressor, and seven items for collaborative activities – Type IV stressor.

The scale was translated into Japanese and checked by two professionals for readability and accuracy. Each question consisted of a Likert-scale measuring the level of perceived mental effort: 1 – Not at all, 2 – Slightly, 3 – Moderately, 4 – Mostly, 5 – Very much in Study 1. In Study 2 the scale was expanded to seven (1 – Not at all ------ 7 – Very much) in order to avoid skew of data.

Cognitive loads. Both Study 1 and 2 employed three questions in order to measure three kinds of cognitive load (intrinsic, extraneous, and germane cognitive load) as were developed by Cierniak, Scheiter, and Gerjets (2009). The three questions below were suitable for the purpose of the dissertation. The question for intrinsic cognitive load was about the difficulty in the learning process and composition task

using worked example. The question for extraneous cognitive load focused on the difficulty of learning content itself that directly detects the excessive amount of mental effort resulted from a difficult task. The question for germane cognitive load was interested in the motivational aspect of mental effort that could be distinguished from the two other cognitive loads. Cierniak et al. (2009) report the three questions to be endorsed for the measurement of three distinctive cognitive loads.

Intrinsic cognitive load: "How difficult was the learning content for you?"
Extraneous cognitive load: "How difficult was it for you to learn?"
Germane cognitive load: "How much did you concentrate during learning?"

The combination of simple questions and rating scale for cognitive load measurement is a common and reliable way of gauging mental effort and is known as being one of the best ways, considering effectiveness and cost (Brünken, Plass, & Leutner, 2003). Kalyuga, Chandler, and Sweller (1998) recommends to employ simple questions and rating scale and found them to be valid in respect to cognitive loads, despite the risk of the objective validity inherent in the use of such subjective methods. Each question consists of a Likert-scale rating to measure the level of perceived mental effort. Study 1 used a five-point scale for the measurement of cognitive load replicating Cierniak et al. (2009). However, the results on cognitive loads were rather deflected around the middle, thus the researcher adopted a nine-point scale that was widely used and verified in various studies (e.g. Stark, Mandl, Gruber, & Renkl, 2002; van Merriënboer, Schuurman, de Croock, & Paas, 2002) in order to gain more precise scaling measures. All questions were translated into Japanese and checked by three professionals for accuracy and clarity.

Collaborative composition quality. Two frameworks were adopted to measure the quality of the collaborative composition task in both Study 1 and 2: composition quality ratings by *ESL composition profile* (Jacobs, Zinkgraf, Wormuth, Hartfiel, & Hughey, 1981) and quantitative analysis by number of T-units and number of words in the composition. A combination of descriptive and quantitative methods was employed for the measurement of collaborative composition performance. The ESL composition profile is widely implemented as a performance measurement by researchers of applied linguistics around the world. Since it was developed for measuring the written performance of English compositions made by non-native speakers, many Japanese EFL researchers use it for the performance criteria (Yamanishi, 2004).

The ESL composition profile contains five criteria: (a) Content, (b) Organization, (c) Vocabulary, (d) Language use, and (e) Mechanics, to evaluate the quality of written products (see Appendix E). Each of these segments has its own evaluation criteria and designated points. Taking the Content segment as an example, it has "Excellent to very good – 30-27 points", "Good to average – 26-22 points", "Fair to poor – 21-17 points" and "Very poor – 16-13 points".

Second, the number of T-units per sentence and the number of words in the collaborative composition were also used to measure the complexity of composition performance. A T-unit (Hunt, 1996) is a phrasal unit for constructing a sentence. The numbers of T-units can indicate the complexity of the sentence.

Procedure

Study 1. Study 1 was conducted in the spring semester, June to August 2012.

It took place under a blended learning condition, which combined a 15-lecture face-to-face credited course, and Moodle was used as the platform for online communication both on and off-campus, in which all the information, directions and instructions about the online collaboration were included. In addition to Moodle, Google Drive was used during off-campus collaboration. The "Forum" function of the Moodle system was used for making personal relationships among members, brainstorming and idea gathering in the initial stage of collaboration. Figure 3-1 describes the design and procedure of the experiment.

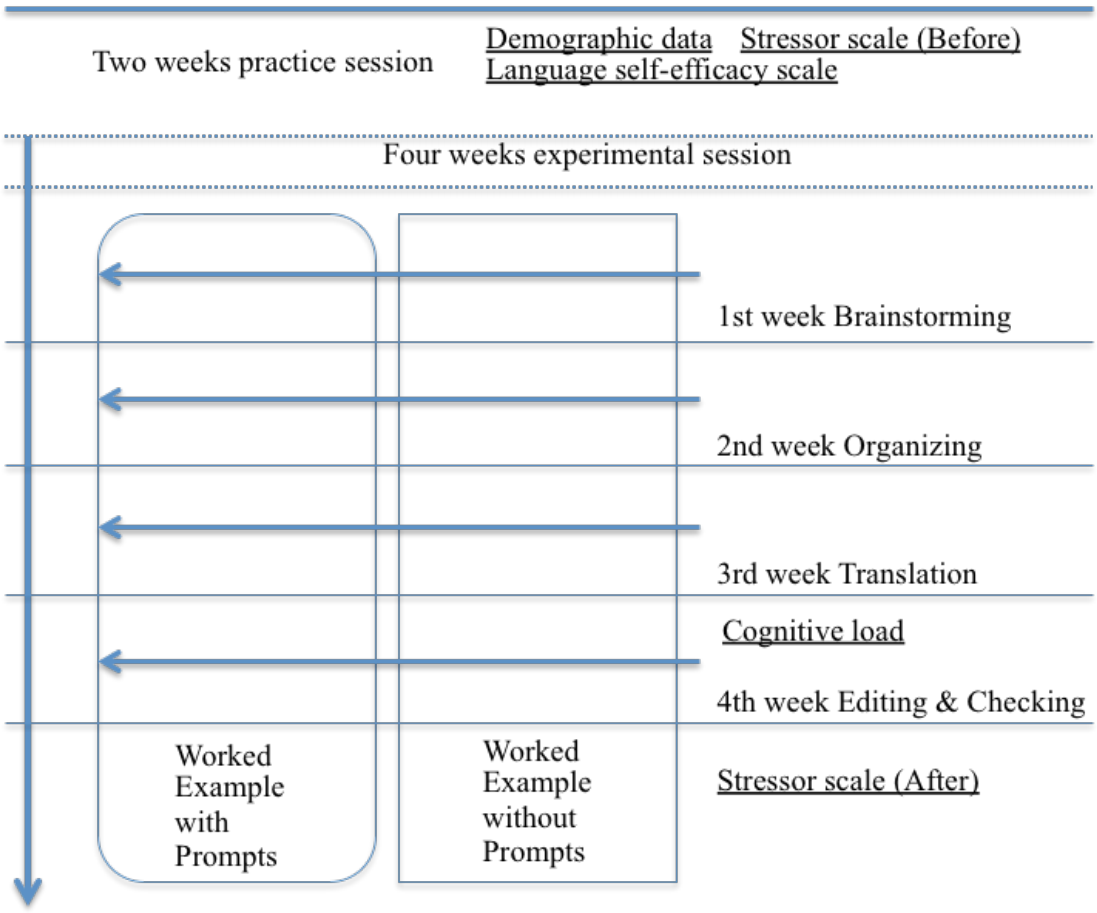

Figure 3-1. The six weeks procedure of Study 1

Practice session. Study 1 was six-weeks long and had two phases. A two-week practice session preceded a four-week experimental session. The practice session started on the 9th week of the course, after six lectures on English grammar for collaborative composition task using original lesson material.

A stressors scale for Before (Jung et al., 2012; Appendix A), language self-efficacy scale (Shaw, 2007; Appendix B), and demographic questions (e.g., sex, age, experience in online learning, preference on collaborative learning, preference on English) were completed by the participants using the paper version, after which they received a copy of textbook and sat lectures by the examiner in the first week of the practice session. The textbook included two short example compositions. A guideline for collaboration, directing them to observe the rules and manners of online collaboration, was provided orally.

The instructor guided participants to follow four phases (Brainstorming phase, Organizing phase, Composition & Translation phase and Editing & Checking phase) to complete the collaborative composition task, as exemplified in the textbook during the practice session. The participants completed one short collaborative composition project during the practice session as homework. In the second week of the practice session, the experimenter reminded participants of the four steps of collaborative composition and structure of English composition in the face-to-face class using the textbook.

Experimental session. After the practice session, the experimental session took place from weeks 11 to 14. During the experimental session, participants received instruction in basic composition process and composition procedures during the face-to-face lectures using a textbook. Although the participants were allowed to talk about the collaborative projects with their peers during face-to-face classes, they were encouraged to work on the composition strictly by collaborative effort online.

In the second week (Organizing phase) of the experimental session, groups received instructional supports. The experimenter posted one type of worked example on the collaborative workspace for each group to follow. The worked example group received a worked example without prompts. The worked example with prompts group received prompts in addition to the same worked example.

In the third week (Composition & Translation phase) of the experimental session, the experimenter carried out a cognitive load survey. This phase was the most effort consuming during the experimental session. Three cognitive loads were measured. The survey was conducted on the Moodle system using the feedback function.

In the final week (Editing & Checking phase) of the experimental session, learners were directed to finish collaborative work within the week. Immediately after the final week, the paper version of the stressor scale (Jung et al., 2012) and the language self-efficacy scale were completed for the second time. This allowed the experimenter to investigate the impact of instructional strategies (two types of worked examples) upon stressors during completion of the collaborative tasks. All participants answered the 32-item questionnaire in five days.

Study 2. Study 2 was conducted in the fall semester, October to December 2012. Similar to Study 1, online collaboration was combined with face-to-face learning in one course. Figure 3-2 shows the procedure of Study 2.

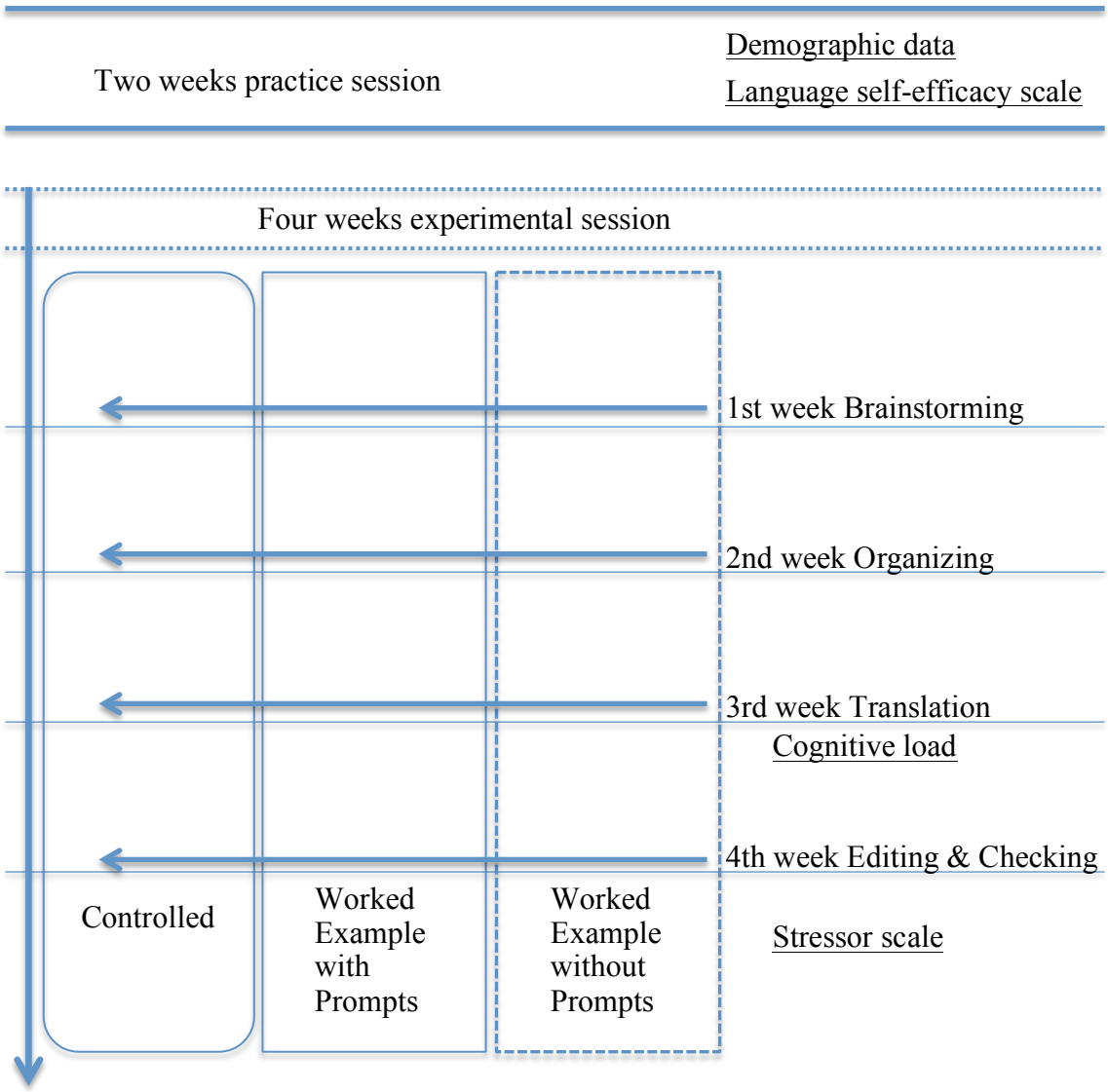

Figure 3-2. The six weeks procedure of Study 2

Practice session. Same as Study 1, the experiment was also two-phased and six-weeks long. A two-week practice session preceded a four-week experimental session. The instructor made sure to guide participants by taking four phases (Brainstorming phase, Organizing phase, Composition & Translation phase, and Editing & Checking phase), as exemplified in the textbook during the practice session.

Experimental session. Study 2 followed the procedure of Study 1. The major difference was the number of measurements of stressors. Study 2 measured the stressors one time. In the third week of the experimental session, questions on cognitive loads were surveyed via the "Form" function of Google Drive. The researcher made a form using the system, and sent out a message containing it to the registered participants. As soon as the participants answered the questionnaire, the system sent results back to the experimenter automatically. The experimenter asked participants to provide their real names and partners' names, which allowed the experimenter to confirm the integrity and genuineness of the answers. The survey was received from all participants within four days.

Immediately after the final week of the experimental session, a stressor scale (Jung et al., 2012) was completed. It was also sent to each participant using the Google Drive Form. All the participants answered the 32-item survey in five days.

Data Collection and Analysis

Data collection. In Study 1, the language self-efficacy scale and stressor scale (Before) was administered during the 1st week of experimental session. In both Study 1 and 2, then stressor scale (After) was administered and three questions on cognitive loads were conducted.

As for perceived stressors, in Study 1, the stressor scale was distributed in the first week (Before) and final week (After) of the experiment, whereas in Study 2, stressor scale was distributed once during the final week of experimental session. The Study 1 informed that a change in stressors took place before and after the treatment, therefore, Study 2 was designed to administer stressor scale once, after the experiment.

Study 1 utilized the paper and pencil version for stressor scales, whereas Study 2 employed Google Drive Form.

As for cognitive loads, in Study 1, cognitive loads were measured in the third week of the experimental session via Moodle. The participants who were absent from the class were asked to answer the survey later. In Study 2, Google Drive Form was used the same three questions for each cognitive load in the third week of the experimental session.

As for collaborative composition, during four weeks of experimental session, the students were assigned to compose a letter. The content was an essay aimed at persuading high school students to choose Sapporo as a location for their college education. After the experimental session of four weeks, the participants were asked to submit the both hard copy and electronic versions in Microsoft Word to the instructor.

Data analysis. The data gathered from both studies were analyzed in a similar way except for the two records of stressors in Study 1 as indicated above. The preliminary analysis on numeric data was computed to obtain summary statistics using computer software, namely: Microsoft Excel and IBM SPSS Statistics version 21.

In order to answer Research Question one, both Study 1 and 2 examined the effect of two treatments on stressors I – IV (Jung et al., 2012) with considering other factors. In Study 1, three-way analyses of variance (ANOVA) with the treatment factor (worked example with prompts group and worked example without prompts group), LSE factor (Low group and High group) and Before-After factor (Before condition and After condition) were carried out. This analysis was conducted for each of the four stressors (I – IV). In Study 2, two-way ANOVA with the treatment factor (Control group, worked example with prompts group and worked example without prompts

group) and the LSE factor (Low group and High group) was conducted for each of the four stressors (I – IV).

In order to answer Research Question two, both Study 1 and 2 examined the effect of the treatments on three kinds of cognitive load (intrinsic, extraneous and germane) with considering other factors. For this purpose, in Study1, two-way ANOVA with the treatments (worked example with prompts group and worked example without prompts group) and LSE factor (Low group and High group) was carried out. Moreover, in Study 2, two-way ANOVA with the treatment factor (control group, worked example with prompts group and worked example without prompts group) and the LSE factor (Low group and High group) was conducted. These analyses were conducted for each of cognitive load.

In order to answer Research Question three, the performance of groups was analyzed by the alternation of the treatments. Study1, t-tests was conducted to examine the difference of the treatments (worked example with prompts group and worked example without prompts group) on each composition performance criteria. Moreover, in Study 2 employed a two-way ANOVA with the treatment factor (control group, worked example with prompts group and worked example without prompts group) on each composition performance criteria.

To analyze the quality of collaborative composition, five evaluation criteria, (a) Content, (b) Organization, (c) Vocabulary, (d) Language use, and (e) Mechanics were evaluated following the ESL composition profile by Jacobs et al. (1981). Each criterion has four levels (Excellent to Very Good, Good to Average, Fair to Poor and Very Poor) and the level owns a range of scores according to the quality of collaborative composition (cf. Excellent to Very Good – 30-27, Poor and Very Poor – 16-13). The highest score is 100 and the lowest is 34.

Two raters specializing in linguistics and education evaluated together each composition based on the criteria listed by the ESL composition profile. The raters evaluated the levels evaluation first by themselves. Study 1 obtained 80.0 % of agreement the evaluation at this stage on the five criteria for quality of collaborative composition and Study 2 obtained 78.2% of agreement. When raters disagreed, they discussed appropriate level for evaluation and decided on the levels. Then, the raters discussed to decide the scores of criteria according the range of points indicated in the profile.

In addition to the composition quality ratings, two other measurements were used to measure composition complexity: the number of T-units per sentence and the number of words in the composition. The same two raters counted all the compositions for the numbers of t-unit manually (inter-rater reliability, Study 1: 95.6 %; Study 2: 97.2%). When there is a disagreement, they discussed and agreed on the numbers. The number of words was counted by the "word count" function of Microsoft Word 2011.

CHAPTER 4 RESULTS

Results of Study 1

Study 1 was administered to answer three Research Questions by employing the methodologies described in Chapter 3. In this section, first, t-tests on each stressor of pre-experiment status were indicated to check the homogeneity among the treatment groups. Second, the results of three-way ANOVA that examined four stressors as dependent variables were shown. Third, the results of two-way ANOVA examined three kinds of cognitive loads as dependent variables were indicated. Finally, the results of t-tests on the difference of the treatments by the performance of groups were presented.

The table 4-1 shows the descriptive statistics of the measured scores and coefficient of reliability for stressors and cognitive loads as dependent variables, LSE that is quantitative variable. LSE scores were gathered and added. The descriptive statistics for LSE were as follows: minimum score (*Min*) = 310, maximum score (*Max*) = 3420, mean (*M*) = 1775.17, standard deviation (*SD*) = 809.82, mean standard error (*Mse*) = 134.97. Coefficient of reliability for reading skill self-efficacy was α = .92; reading tasks self-efficacy was α = .97; writing skill self-efficacy was α = .95; and writing task self-efficacy was α = .95. LSE was divided by the method explained in Chapter 3 based on the mean values and standard deviation.

Table 4-1
The Descriptive Statistics of the Measured Scores and Coefficient of Reliability for Stressors and Cognitive Loads

	Before Condition (*n*=36)					After Condition (*n*=36)					
	M	*SD*	*Mse*	*Min*	*Max*	*M*	*SD*	*Mse*	*Min*	*Max*	*α*
Type I	43.08	8.22	1.37	19	60	39.75	1.41	8.47	24	57	0.92
Type II	19.67	5.17	0.86	7	31	19.22	0.71	4.27	8	28	0.97
Type III	14.03	4.21	0.70	6	25	13.42	0.88	5.29	6	26	0.95
Type IV	20.39	4.94	0.82	7	30	22.14	0.75	4.47	13	31	0.95
Intrinsic CL						3.81	0.14	0.86	1	5	
Extraneous CL						3.64	0.14	0.87	1	5	
Germane CL						3.92	0.13	0.77	2	5	

Note. Cognitive loads (CL) were measured one time during experimental period and gauzed by one question. Therefore, coefficient of reliability is not indicated.

Worked examples and stress. Preliminary t-tests examined if results were affected by the precondition for influence of stressors. As shown in the Table 4-2, the mean values of stressors (I –IV) before the experiment of Study 1 and standard deviations of the groups are calculated. A t-test of equality revealed no significant group differences among the pre-evaluation groups: t (34) = 0.18, p = .86, d = 0.06 for Type I stressor, t (34) = -0.32, p = .75, d = 0.11 for Type II stressor, t (34) = -1.24, p = .22, d = 0.41 for Type III stressor, t (34) = -0.13, p = .90, d = 0.04 for Type IV stressor. Hence, the experimental groups were comparable with respect to level of influence from stressors.

Table 4-2
The Means and SDs of Pretest Stressor Measurements and Results of t-test

	WE with prompts (*n*=18)			WE without prompts (*n*=18)			Result of *t*-test			
	M	*SD*	*Mse*	*M*	*SD*	*Mse*	*t*	*df*	*p-value*	*Effect size (d)*
Type I	42.83	9.03	2.13	43.33	7.59	1.79	0.18	34	0.86	0.06
Type II	19.94	5.64	1.33	19.39	4.79	1.13	0.32	34	0.75	0.11
Type III	14.89	4.06	0.96	13.17	4.29	1.01	1.24	34	0.22	0.41
Type IV	20.50	5.98	1.41	20.28	3.79	0.89	0.13	34	0.90	0.04

Regarding Research Question 1, the examination of the interactions of the three factors could elucidate inter-relation of factors. Therefore, a three-way ANOVA (not-repeated / repeated combined measures) on the mean values of Type I stressor comparing all the pairs of three variables was conducted.

Effects on Type I stressor. A three-way ANOVA on the mean values of Type I stressor revealed a significant secondary interaction between a Before-After factor, a treatment factor, and a LSE factor, F (1, 27) = 4.66, p = .040, *partial* η^2 = .147. Table 4-3 shows the mean values and *SDs* of Type I stressor in each group.

Table 4-3
The Means and SDs of Type I Stressor Measurements in Each Group

			Before		After	
Treatment	LSE	*n*	*M*	*SD*	*M*	*SD*
WE with prompts	Low LSE	8	48.25	5.80	40.25	6.80
	High LSE	6	39.44	9.00	36.67	5.96
WE without prompts	Low LSE	8	46.13	6.73	47.75	6.52
	High LSE	9	43.67	6.92	42.17	7.83

Since a secondary interaction between three factors (the Before-After factor, the treatment factor, and the LSE factor) was found, a test of simple interactions among the three factors was conducted. In this test, the three factors were examined in turn, to examine the interaction with the levels of other two factors.

At first, simple interactions between the Before-After factor and the treatment factor under each LSE group (Low and High) were examined. While the test on the Low LSE group revealed a statistically significant simple interaction between the Before-After factor and the treatment factor, F (1, 14) = 15.538, p = .001, *partial* η^2 = .526, the test on the High LSE group detected no simple interaction, F (1, 13) = 0.177, p = .681, *partial* η^2 = .013.

Since a simple interaction on low the LSE group was found, a test of simple main effect of the treatment factor on the Before-After factor was examined. In regard to the Low LSE group, the simple main effect of the treatment factor on the Before condition was not statistically significant, F (1, 14) = 0.458, p = .510, *partial* η^2 = .032, while the simple main effect of the treatment factor on the After condition, F (1, 14) = 5.072, p = .041, *partial* η^2 = .266, was statistically significant. When compared with the mean value of Type I stressor in After condition, the WE with prompts group scored significantly lower than that of the WE without prompts group. Moreover, under the grouping of Low LSE, the test of simple main effect of the Before-After factor on the treatment factor condition was examined. For the WE without prompts group, the simple main effect of the Before-After factor was not significant, F (1, 14) = 0.886, p = .363, *partial* η^2 = .060; however the simple main effect of the Before-After factor on the WE with prompts group, F (1, 14) = 21.469, p = .001, *partial* η^2 = .605, was statistically significant. The mean value of the Type I stressor of the WE with prompts group scored significantly lower in the After condition than in the Before condition.

For the High LSE group, the main effect of a treatment factor, F (1, 13) = 1.751, p = .209, *partial* η^2 = .119, and the main effect of the Before-After factor, F (1, 13) = 1.987, p = .182, *partial* η^2 = .133 were not statistically significant.

Figure 4-1 illustrates the differences in mean values of the Type I stressor for each condition of the Before-After factor and the treatment factor under LSE groups (Low and High).

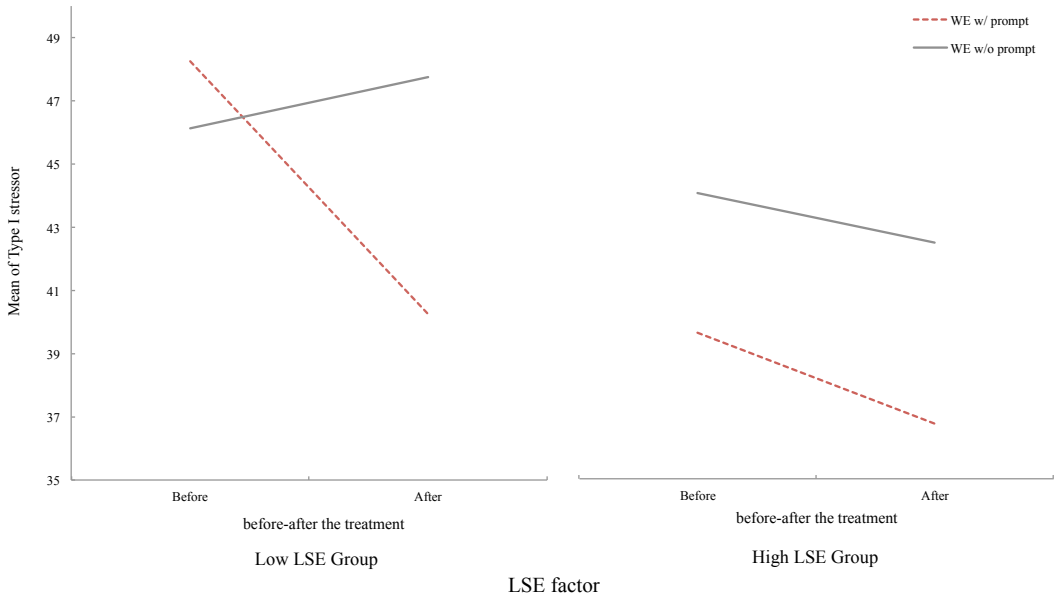

Figure 4-1. The summary of mean values of Low and High LSE groups

Second, simple interactions between the treatment factor and the LSE factor under each Before-After condition were examined. In regard to the Before condition, there was no interaction, F (1, 27) = 1.431, p = .242, *partial* η^2 = .050, but the main effect of the LSE factor was found, F (1, 27) = 4.508, p = .043, *partial* η^2 = .143, whereas no main effect for the treatment factor, F (1, 27) = 0.156, p = .696, *partial* η^2 = .006, was found. The mean value of Type I stressor scored statistically lower in the High LSE group than that of the Low LSE group in the Before condition.

As for the After condition, there was no statistically significant interaction, F (1, 27) = 0.169, p = .684, *partial* η^2 = .006, but the main effect of the treatment factor was found, F (1, 27) = 7.134, p = .013, *partial* η^2 = .209, while the main effect of the LSE factor was not found, F(1, 27) = 3.547, p = .070, *partial* η^2 = .116. The mean value of Type I stressor scored statistically lower in the WE with prompts group than that of the WE without prompts group. Figure 4-2 illustrates the differences in mean values of the Type I stressor for each treatment factor and each LSE factor under the Before-After condition.

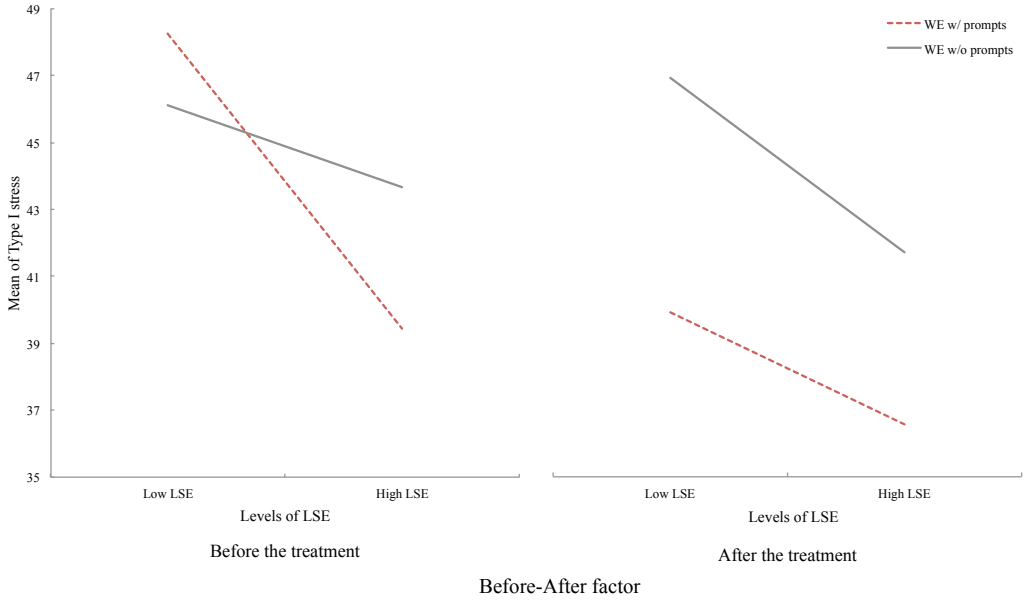

Figure 4-2. The summary of mean values of Before and After conditions

Third, simple interactions between the Before-After factor and the LSE factor under each treatment condition (WE with prompts and WE without prompts) were examined. The summary of each group is shown in Figure 4-3. Both tests on the WE without prompts group, F (1, 12) = 1.901, p = .193, *partial* η^2 = .137, and the WE with prompts group, F (1, 15) = 3.130, p = .097, *partial* η^2 = .173, found no interaction. There was, however, a main effect of the Before-After factor on the WE with prompts group, F (1, 15) = 13.331, p = .002, *partial* η^2 = .471. The mean value of Type I stressor scored statistically lower in the After condition than that of the Before condition.

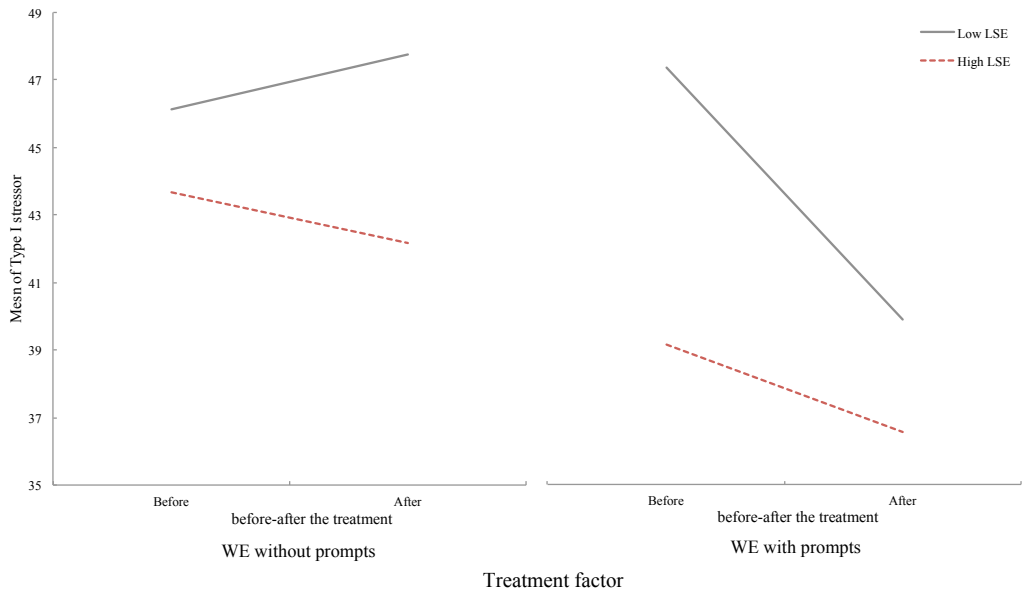

Figure 4-3. The summary of mean values of WE without prompts and WE with prompts groups

Effects on Type II stressor. A three-way ANOVA (repeated / not-repeated combined measures) on the mean values of Type II stressor comparing all the pairs of three variables was conducted. The summary of Type II stressor is shown in Table 4-4.

Table 4-4
The Means and SDs of Type II Stressor Measurements in Each Group

Treatment			Before		After	
	LSE	*n*	*M*	*SD*	*M*	*SD*
WE with prompts	Low LSE	8	19.88	2.03	17.00	5.29
	High LSE	6	21.44	6.37	19.78	4.06
WE without prompts	Low LSE	8	19.75	4.71	21.13	4.32
	High LSE	9	20.67	4.13	20.17	2.79

The tests on the mean values of Type II stressor on three variables did not reveal any significant secondary interactions, F (1, 27) = .865, p = .361, *partial* η^2 = .031. Moreover, primary interactions of the Before-After factor * the treatment factor, the Before-After factor * the LSE factor, the treatment factor * the LSE factor were not statistically significant, either. Also, there was no statistically significant main effect on

any factor, either.

Effects on Type III stressor. A three-way ANOVA (repeated / not-repeated combined measures) on the mean values of Type III stressor comparing all the pairs of three variables was conducted. The summary of Type III stressor is shown in Table 4-5.

Table 4-5
The Means and SDs of Type III Stressor Measurements in Each Group

			Before		After	
Treatment	LSE	*n*	*M*	*SD*	*M*	*SD*
WE with prompts	Low LSE	8	15.75	4.03	12.88	6.24
	High LSE	6	15.11	3.18	15.33	3.00
WE without prompts	Low LSE	8	14.00	5.24	15.63	6.84
	High LSE	9	14.00	3.52	12.17	4.40

The tests on the mean values of Type III stressor on three variables did not reveal any significant secondary interactions, $F\ (1, 27) = 3.809$, $p = .061$, *partial* $\eta^2 = .124$. Moreover, primary interactions the Before-After factor * the treatment factor, the Before-After factor * the LSE factor, the treatment factor * the LSE factor were not statistically significant. Also, there was no statistically significant main effect on any factor, either. Therefore, no statistically significant difference was scored between the factors.

Effects on Type IV stressor. A three-way ANOVA (repeated / not-repeated combined measures) on the mean values of Type IV stressor comparing all the pairs of the three variables above was conducted. The summary of Type IV stressor is shown in Table 4-6.

Table 4-6
The Means and SDs of Type IV Stressor Measurements in Each Group

Treatment			Before		After	
	LSE	*n*	*M*	*SD*	*M*	*SD*
WE with prompts	Low LSE	8	22.63	5.13	22.38	4.81
	High LSE	6	20.11	5.06	19.89	3.72
WE without prompts	Low LSE	8	19.75	4.13	26.00	3.51
	High LSE	9	20.67	3.88	21.50	4.09

As a result of three-way ANOVA on the mean values of Type IV stressor, no secondary interaction was found, F (1, 27) = 2.387, p = .134, *partial* η^2 = .081. However, a primary interaction between the Before-After factor and the treatment factor was found significant, F (1, 27) = 4.596, p = .041, *partial* η^2 = .145, while a primary interaction between the Before-After factor and the LSE factor, F (1, 27) = 2.338, p = .138, *partial* η^2 = .080, and between the treatment factor and the LSE factor, F (1, 27) = 0.072, p = .790, *partial* η^2 = .003, were not statistically significant. Thus, a test of simple main effect on the Before-After factor and the treatment factor in an interaction between the Before-After factor and the LSE factor was conducted.

The simple main effect of the Before-After factor on the WE with prompts group was not statistically significant, F (1, 27) = 0.040, p = .843, *partial* η^2 = .001, while the Before-After factor on the WE without prompts group was statistically significant, F (1, 27) = 7.310, p = .012, *partial* η^2 = .213. Figure 4-4 describes the summary. The mean value of Type IV stressor of the WE without prompts group for the After condition was higher than that of the Before condition. However, simple main effect of the treatment factor for the Before condition, F (1, 27) = 0.472, p = .498, *partial* η^2 = .017, was not statistically significant; nevertheless simple main effect of the treatment factor for the After condition, F (1, 27) = 3.171, p = .086, *partial* η^2 = .105, showed a statistically significant tendency.

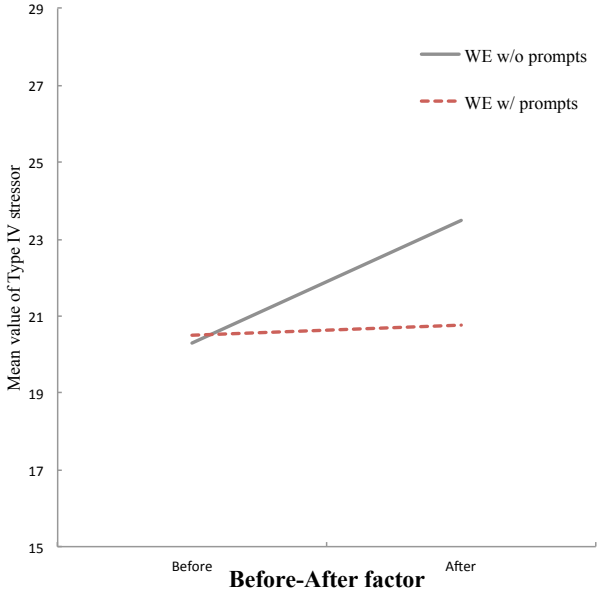

Figure 4-4. The summary of mean values of Type IV stressor

Worked examples and cognitive loads. Regarding Research Question two, two-way ANOVA tests were used to examine the difference in three kinds of cognitive load measurements (intrinsic, extraneous, and germane) and the effects of two between-subjects factors, the treatments (WE with prompts and WE without prompts) and the LSE (Low and High).

Effects on intrinsic cognitive load. First, a two-way ANOVA was conducted to examine the effect of a treatment factor and a LSE factor on intrinsic cognitive load. The mean values and *SDs* of intrinsic cognitive load are shown in Table 4-7.

Table 4-7
The Means and SDs of Intrinsic Cognitive Load Measurements in Each Group

Treatment	LSE	*n*	*M*	*SD*
WE with prompts	Low LSE	8	4.38	0.74
	High LSE	6	3.44	0.53
WE without prompts	Low LSE	8	3.88	0.64
	High LSE	9	4.00	0.63

The test found a significant interaction between the treatment factor and the LSE factor, F (1, 27) = 5.195, p = .031, *partial* η^2 = .161. Therefore, tests of simple main effect were conducted. The simple main effect of the LSE factor on the WE with prompts group was statistically significant, F (1, 27) = 9.025, p = .006, *partial* η^2 = .251, while the simple main effect of the LSE factor on the WE without prompts group was not statistically significant, F (1, 27) = 0.132, p = .719, *partial* η^2 = .005. Figure 4-5 describes the summary. The mean value of the intrinsic cognitive load of the WE with prompts group for the High LSE group was lower than it was for the Low LSE group. However, the simple main effect of the treatment factor under both the Low LSE group, F (1, 27) = 2.461, p = .128, *partial* η^2 = .084, and the High LSE group, F (1, 27) = 2.734, p = .110, *partial* η^2 = .092, were not statistically significant.

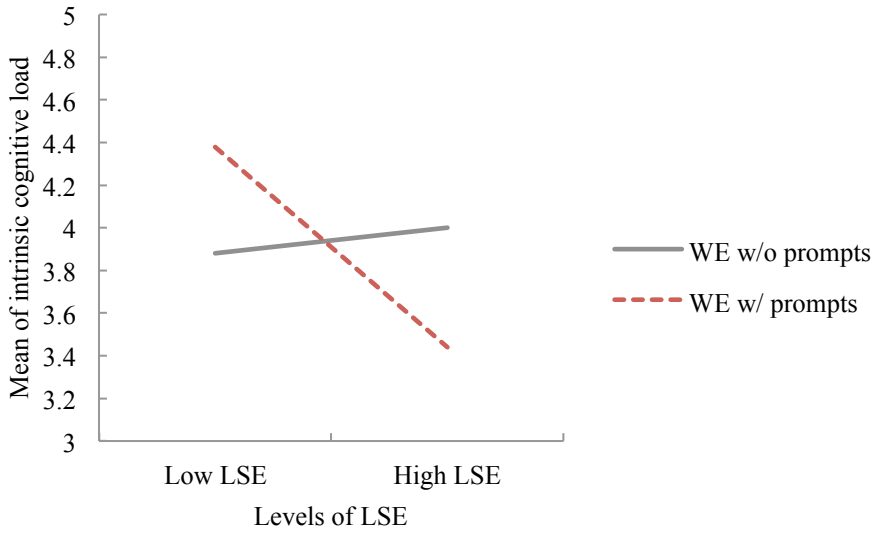

Figure 4-5. The summary of mean values of intrinsic cognitive load

Effects on extraneous cognitive load. Another two-way ANOVA on the mean values of extraneous cognitive load was conducted. The summary of extraneous cognitive load is shown in Table 4-8.

Table 4-8

The Means and SDs of Extraneous Cognitive Load Measurements in Each Group

Treatment	LSE	*n*	*M*	*SD*
WE with prompts	Low LSE	8	4.00	1.07
	High LSE	6	3.44	0.53
WE without prompts	Low LSE	8	3.75	0.71
	High LSE	9	3.83	0.41

The test found no interaction between the treatment factor and the LSE factor, F (1, 27) = 1.435, p = 0.241, *partial* η^2 = .050, nor main effects of the treatment factor, F (1, 27) = .068, p = 0.797, *partial* η^2 = .003, or the LSE factor, F (1, 27) = .784, p = 0.384, *partial* η^2 = .028. Figure 4-6 describes the summary.

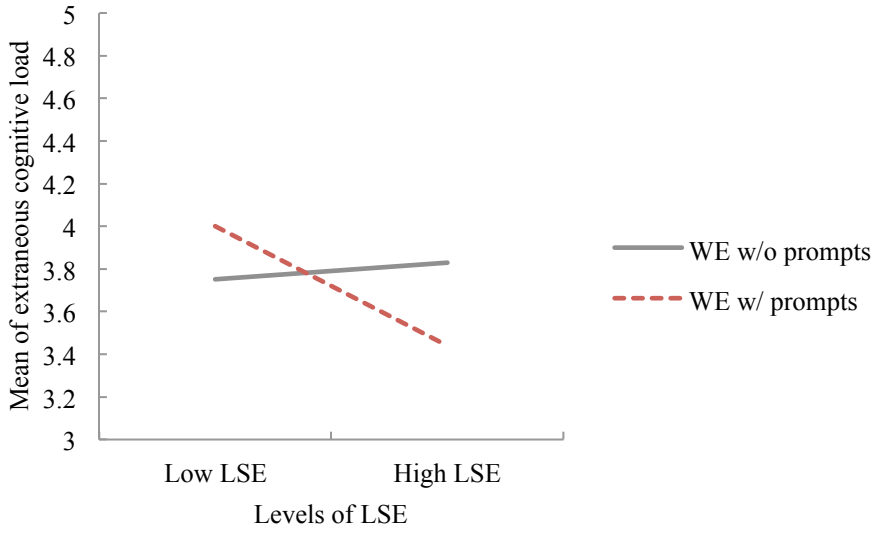

Figure 4-6. The summary of mean values of extraneous cognitive load

Effects on germane cognitive load. For germane cognitive load, two-way ANOVA on the mean values of germane cognitive load was conducted. The summary of germane cognitive load is shown in Table 4-9.

Table 4-9

The Means and SDs of Germane Cognitive Load Measurements in Each Group

Treatment	LSE	*n*	*M*	*SD*
WE with prompts	Low LSE	8	3.75	0.89
	High LSE	6	4.11	0.33
WE without prompts	Low LSE	8	3.38	0.92
	High LSE	9	4.33	0.52

The test found no interaction among the treatment and the LSE factor, F (1, 27) = 1.342, p = 0.257, *partial* η^2 = .047; however, the test found a main effect of the LSE factor, F (1, 27) = 6.550, p = .016, *partial* η^2 = .195, while no main effect of the treatment factor, F (1, 27) = .088, p = 0.716, *partial* η^2 = .003. The mean value of the germane cognitive load of the High LSE group scored significantly higher than that of the Low LSE group. Figure 4-7 describes the summary.

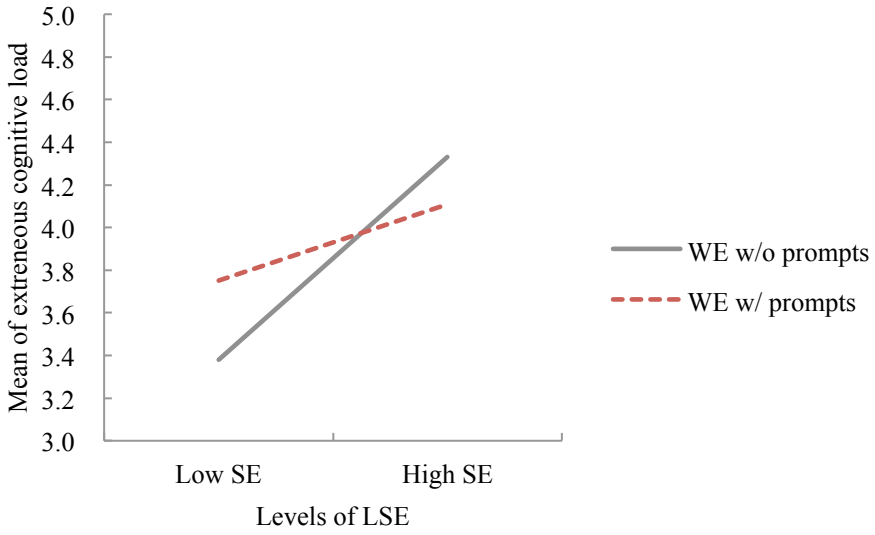

Figure 4-7. The summary of mean values of germane cognitive load

Worked examples and performance in collaborative composition. Regarding Research Question three, first, a two tailed t-tests for comparing mean values

of each collaborative composition performance criteria on the effect of the treatment groups (WE with prompts and WE without prompts) were conducted. A descriptive statistics of quality of collaborative composition is showed in Table 4-10.

Table 4-10
The Comparison of Means by the Treatment Difference of Collaborative Composition Performance and Results of t-test

	WE without prompts (*n*=6)			WE with prompts (*n*=6)			Result of *t*-test			
	M	*SD*	*Mse*	*M*	*SD*	*Mse*	*t*	*df*	*p-value*	*Effect size (d)*
Contents	21.33	4.59	1.87	20.17	3.25	1.33	0.51	10	0.62	0.29
Organization	14.83	4.07	1.66	15.50	3.27	1.34	0.31	10	0.76	0.18
Vocabulary	15.50	3.56	1.45	15.67	1.75	0.71	0.10	10	0.92	0.06
Language	18.50	4.72	1.93	16.83	3.43	1.40	0.70	10	0.50	0.40
Mechanics	3.67	0.52	0.21	3.67	0.52	0.21	0.00	10	1.00	0.00
GT	73.83	16.63	6.79	71.83	11.82	4.83	0.24	10	0.82	0.14

The results from an independent sample two-tailed t-test revealed no difference between the prompted and without group on Contents, t (10) = 0.51, p = .62, d = 0.29; Organization, t (10) = 0.31, p = .76, d = 0.18; Vocabulary, t (10) = 0.10, p = .92, d = 0.06; Language, t (10) = 0.70, p =.50, d = 0.40; Mechanics, t (10) = 0.00, p = 1.00, d = 0.00. Figure 4-8 describes the summary of t-test.

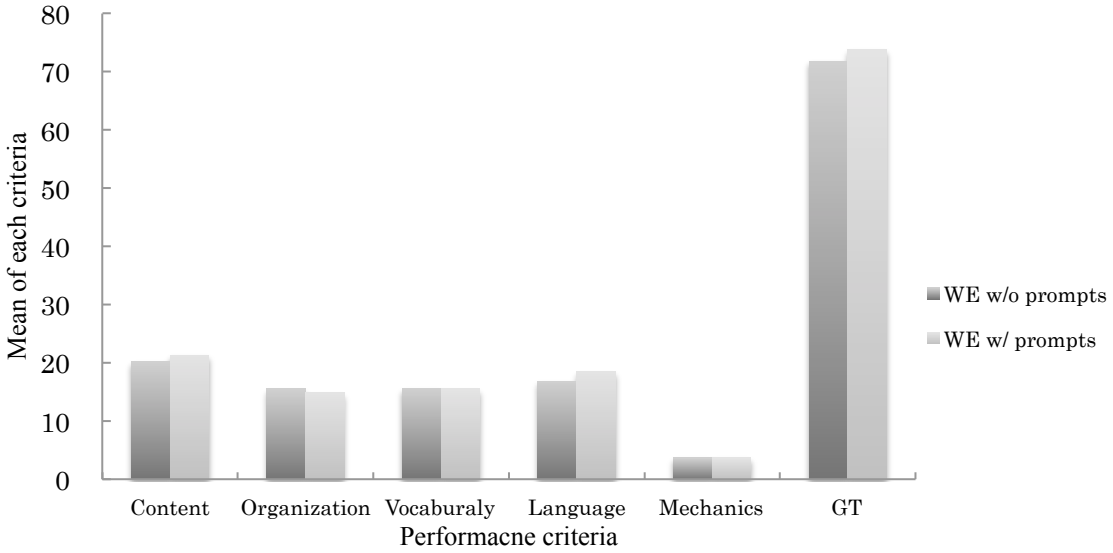

Figure 4-8. The summary of mean values of composition performance of Study 1

Second, a two tailed t-tests for comparing mean values of each index of complexity of composition on the effect of the treatment groups (WE with prompts and WE without prompts) were conducted. A descriptive statistics of index of complexity of composition and results of t-test is showed in Table 4-11.

Table 4-11
The Comparison of Means by the Treatment Difference on Complexity of Composition and Results of t-test

	WE without prompts (*n*=6)			WE with prompts (*n*=6)			Result of *t*-test			
	M	*SD*	*Mse*	*M*	*SD*	*Mse*	*t*	*df*	*p-value*	*Effect size (d)*
No of words	397.83	175.68	71.72	427.00	85.53	34.92	0.37	10	0.72	0.21
No of sentences	30.50	11.35	4.63	34.83	7.65	3.12	0.78	10	0.46	0.45
T-unit per sentence	1.42	0.20	0.08	1.39	0.07	0.03	0.26	10	0.80	0.20

The result from an independent sample t-test revealed no difference between WE with prompts and WE without prompts groups on numbers of words, t (10) = 0.37, p = .72, d = 0.21; numbers of sentences, t (10) = 0.78, p = .46, d = 0.45; numbers of

T-units per sentence, t (10) = 0.26, p = .80, d = 0.20; two-tailed.

Results of Study 2

Study 2 was an attempt to collect further empirical evidence for a firm understanding of the effects of worked examples with and without prompts in the online collaborative learning environment. In this section, first, the results of two-way ANOVA that examined all types of stressors as dependent variables were expressed. Second, the results of two-way ANOVA examined three kinds of cognitive loads as dependent variables were indicated. Finally, the results of one-way ANOVA on the difference of the treatments by the performance of groups were indicated.

The table 4-12 shows the descriptive statistics of the measured scores and coefficient of reliability for stressors and cognitive loads as dependent variables, LSE that is quantitative variable. LSE scores are gathered and added. The descriptive statistics for LSE are as follows: *min* = 40, *max* = 3280, *M* = 1363.86, *SD* = 775.27, and *Mse* = 95.43. Coefficient of reliability for reading skill self-efficacy was α = .95; reading tasks self-efficacy was α = .97; writing skill self-efficacy was α = .95; and writing task self-efficacy was α = .96. LSE was divided by the method explained in Chapter 3 based on the mean value and standard deviation.

Table 4-12
The Descriptive Statistics of the Measured Scores and Coefficient of Reliability for Stressors and Cognitive Loads

	Stressors and Cognitive Loads of Study 2 (*n*=66)					
	M	*SD*	*Mse*	*Min*	*Max*	α
Type I	63.59	1.02	8.26	45	80	0.95
Type II	29.67	0.72	5.82	17	45	0.97
Type III	19.39	0.87	7.03	6	37	0.95
Type IV	30.39	1.15	9.33	7	48	0.96
Intrinsic CL	7.02	0.18	1.42	2	9	
Extraneous CL	6.03	0.25	2.02	1	9	
Germane CL	6.33	0.20	1.63	2	9	

Note. Cognitive loads (CL) were gauzed by one question. Therefore, coefficient of reliability is not indicated.

Worked examples and stress. Regarding Research Question one, two-way ANOVA on the mean values of Type I – IV stressor comparing a treatment factor and a LSE factor was conducted.

Effects on Type I stressor. A two-way ANOVA was conducted to examine the effect of the treatment factor (Control, WE with and without prompts) and the LSE (High and Low) factor on Type I stressor. Table 4-13 shows the mean values and *SDs* of Type I stressor in each group.

Table 4-13
The Means and SDs of Type I Stressor Measurements in Each Group

	LSE	*n*	*M*	*SD*
Control	Low LSE	9	69.22	5.26
	High LSE	10	66.70	8.17
WE with prompts	Low LSE	8	67.25	5.06
	High LSE	10	56.40	6.04
WE without prompts	Low LSE	10	69.30	5.77
	High LSE	7	56.14	8.40

As the result of a two-way ANOVA, there was a statistically significant interaction between the treatment factor and the LSE factor, F (2, 48) = 3.309, p = .045, *partial* η^2 = .121. Since an interaction was found, tests of simple main effect of the LSE factor for three treatment groups were conducted. The simple main effect of the LSE factor for the control group was not statistically significant, F (1, 48) = 0.705, p = .405, *partial* η^2 = .014, while the simple main effect of the LSE factor for the WE with prompts group, F (1, 48) = 12.236, p = .001, *partial* η^2 = .203, and for the WE without prompts group, F (1, 48) = 16.670, p = .001, *partial* η^2 = .258, were statistically significant. For both WE with and without prompts groups, the mean values of the Type I stressor of the High LSE groups were lower than those of the Low LSE groups.

Accordingly, simple main effects of the treatment factor for each LSE group were conducted. The simple main effect of the treatment factor for the Low LSE group was not statistically significant, F (2, 48) = 0.267, p = .767, *partial* η^2 = .011. For the High LSE group; however, the simple main effect of the treatment was statistically significant, F (2, 48) = 7.975, p = .001, *partial* η^2 = .249.

In the High LSE group, a multiple pairwise comparisons using Bonferroni method revealed that the differences in mean values between the Control * the WE with prompts group and the Control * the WE without prompts group were statistically significant, but the differences in mean values between the WE with prompts group and the WE without prompts group were not statistically significant. Figure 4-9 illustrates the differences in the mean values of the Type I stressor for each treatment factor and each LSE factor.

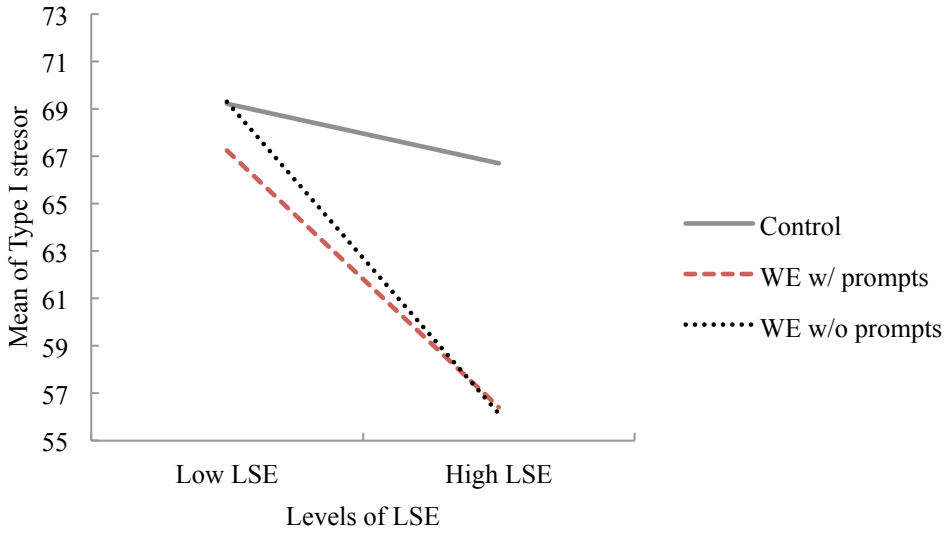

Figure 4-9. The summary of mean values of Type I stressor

Effects on Type II stressor. A two-way ANOVA was conducted to examine the effect of the treatment factor and the LSE factor on Type II stressor. Table 4-14

shows the mean values and *SD*s of Type II stressor for each group.

Table 4-14
The Means and SDs of Type II Stressor Measurements in Each Group

	LSE	*n*	*M*	*SD*
Control	Low LSE	9	31.78	5.19
	High LSE	10	30.30	7.09
WE with prompts	Low LSE	8	30.00	3.96
	High LSE	10	25.30	7.15
WE without prompts	Low LSE	10	32.80	4.02
	High LSE	7	28.00	3.83

As the result of a two-way ANOVA, there was no statistically significant interaction between the treatment factor and the LSE factor, F (2, 48) = 0.532, p = .591, *partial* η^2 = .022. However, there was a main effect of the LSE factor, F (1, 48) = 5.790, p = .020, *partial* η^2 = .108, on Type II stressor; while there was no main effect of the treatment factor, F (2, 48) = 1.904, p = .160, *partial* η^2 = .074. Figure 4-10 illustrates the differences in mean values of the Type II stressor for each treatment factor and each LSE factor.

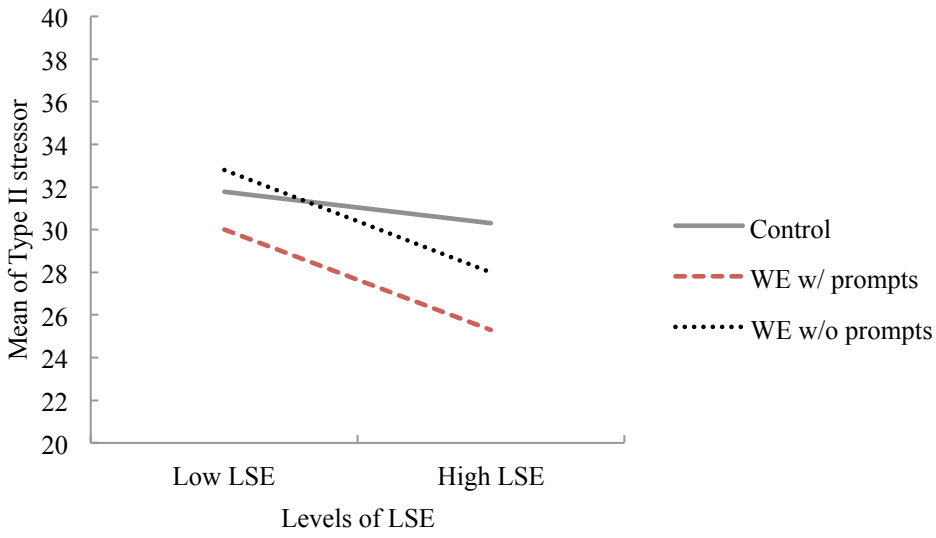

Figure 4-10. The summary of mean values of Type II stressor

Effects on Type III stressor. A two-way ANOVA was conducted to examine the effect of the treatment factor and the LSE factor on Type III stressor. Table 4-15 shows the mean values and *SDs* of Type III stressor for each group.

Table 4-15
The Means and SDs of Type III Stressor Measurements in Each Group

	LSE	*n*	*M*	*SD*
Control	Low LSE	9	19.11	9.94
	High LSE	10	19.50	8.92
WE with prompts	Low LSE	8	18.38	3.38
	High LSE	10	17.00	7.45
WE without prompts	Low LSE	10	19.30	6.65
	High LSE	7	17.14	5.76

As the result of a two-way ANOVA, there was no statistically significant interaction between the treatment factor and the LSE factor, F (2, 48) = 0.137, p = .873, *partial* η^2 = .006. Moreover, there was no main effect of the LSE factor, F (1, 48) = 0.260, p = .612, *partial* η^2 = .005, and the treatment factor, F (2, 48) = 0.224, p = .800, *partial* η^2 = .009. Figure 4-11 illustrates the differences in the mean values of the Type III stressor for each treatment factor and each LSE factor.

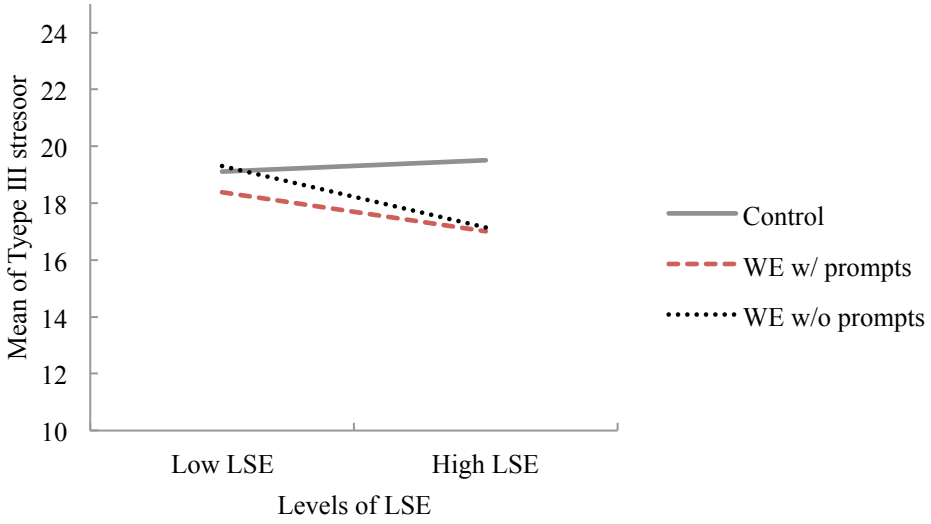

Figure 4-11. The summary of mean values of Type III stressor

Effects on Type IV stressor. A two-way ANOVA was conducted to examine the effect of the treatment and the LSE factor on Type IV stressor. Table 4-16 shows the mean values and *SD*s of Type IV stressor for each group.

Table 4-16
The Means and SDs of Type IV Stressor Measurements in Each Group

	LSE	*n*	*M*	*SD*
Control	Low LSE	9	34.44	2.74
	High LSE	10	29.20	7.84
WE with prompts	Low LSE	8	35.75	7.17
	High LSE	10	23.40	9.30
WE without prompts	Low LSE	10	37.10	6.57
	High LSE	7	23.57	11.00

As the result of a two-way ANOVA, there was no statistically significant interaction between the treatment factor and the LSE factor, F (2, 48) = 1.528, p = .227, *partial* η^2 = .060. However, there was a main effect of the LSE factor, F (1, 48) = 23.891, p = .001, *partial* η^2 = .332, on Type IV stressor; while there was no main effect

of the treatment factor, F (2, 48) = 0.403, p = .671, *partial* η^2 = .017. Figure 4-12 illustrates the differences in the mean values of the Type IV stressor for each treatment factor and each LSE factor.

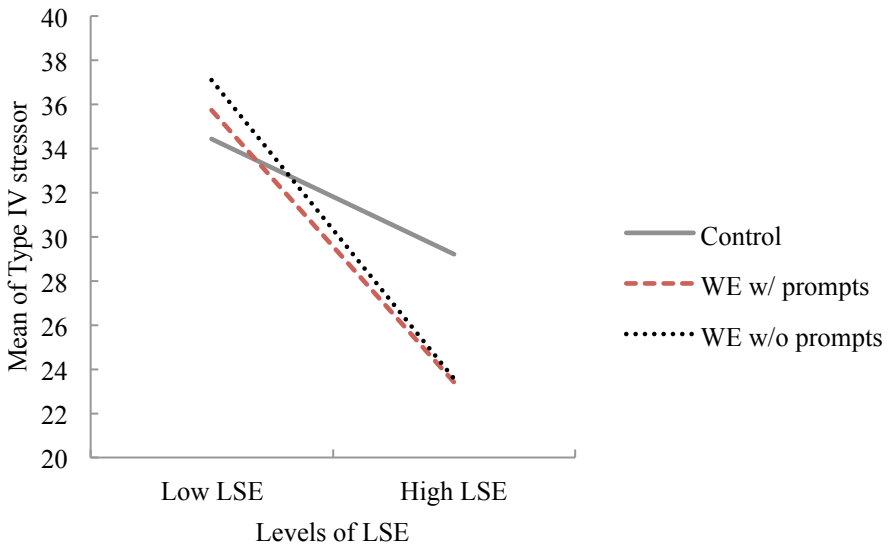

Figure 4-12. The summary of mean values of Type IV stressor

Worked examples and cognitive loads. Regarding Research Question two, two-way ANOVA on the mean values of three kinds of cognitive loads comparing a treatment factor and a LSE factor was conducted.

Effects on intrinsic cognitive load. A two-way ANOVA was conducted to examine the effect of the treatment factor and the LSE factor on intrinsic cognitive load. The descriptive statistics of intrinsic cognitive load for each group are shown in Table 4-17.

Table 4-17

The Means and SDs of Intrinsic Cognitive Load Measurements in Each Group

	LSE	*n*	*M*	*SD*
Control	Low LSE	9	7.44	1.42
	High LSE	10	7.40	0.84
WE with prompts	Low LSE	8	7.75	1.58
	High LSE	10	6.00	1.49
WE without prompts	Low LSE	10	7.80	1.23
	High LSE	7	6.29	1.11

As the result of a two-way ANOVA, there was no statistically significant interaction between the treatment factor and the LSE factor, F (2, 48) = 2.322, p = .109, *partial* η^2 = .088. However, there was a main effect of the LSE factor, F (1, 48) = 9.357, p = .003, *partial* η^2 = .166, while there was no main effect of the treatment factor, F (1, 48) = 0.856, p = .431, *partial* η^2 = .034. The mean value of intrinsic cognitive load was significantly lower in the High LSE group than that of the Low LSE group. Figure 4-13 shows the mean of each value.

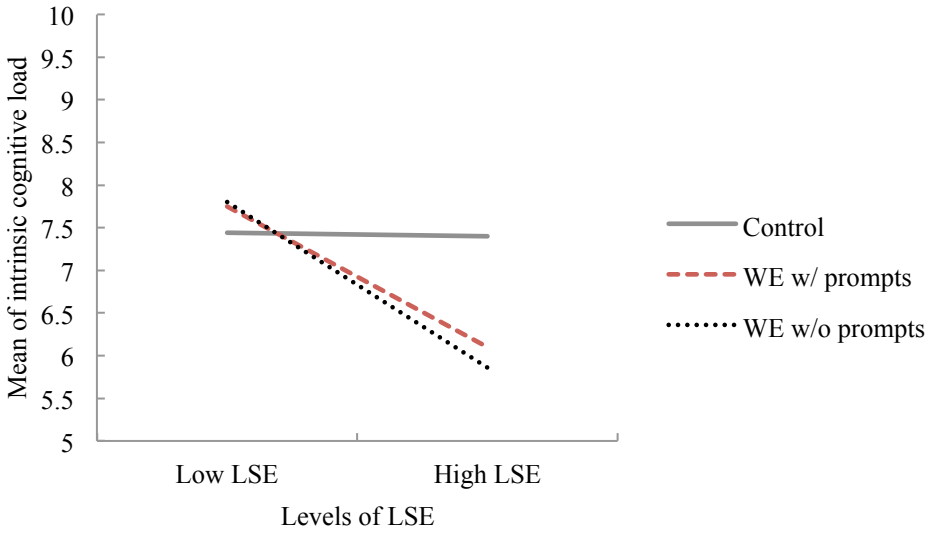

Figure 4-13. The summary of mean values of intrinsic cognitive load

Effects on extraneous cognitive load. A two-way ANOVA was conducted to examine the effect of the treatment factor and the LSE factor on extraneous cognitive load. The descriptive statistics of extraneous cognitive load for each group are shown in Table 4-18.

Table 4-18
The Means and SDs of Extraneous Cognitive Load Measurements in Each Group

	LSE	*n*	*M*	*SD*
Control	Low LSE	9	6.56	1.42
	High LSE	10	6.30	1.57
WE with prompts	Low LSE	8	6.88	1.96
	High LSE	10	5.00	2.26
WE without prompts	Low LSE	10	7.20	2.04
	High LSE	7	5.29	2.50

As the result of a two-way ANOVA, there was no statistically significant interaction between the treatment factor and the LSE factor, F (2, 48) = 1.055, p = .356, *partial* η^2 = .042. However, there was a main effect of the LSE factor, F (1, 48) = 6.211, p = .016, *partial* η^2 = .115, while there was no main effect of the treatment factor, F (1, 48) = 1.118, p = .751, *partial* η^2 = .012. The mean value of extraneous cognitive load was significantly lower in the High LSE group than that of the Low LSE group. Figure 4-14 shows the mean of each value.

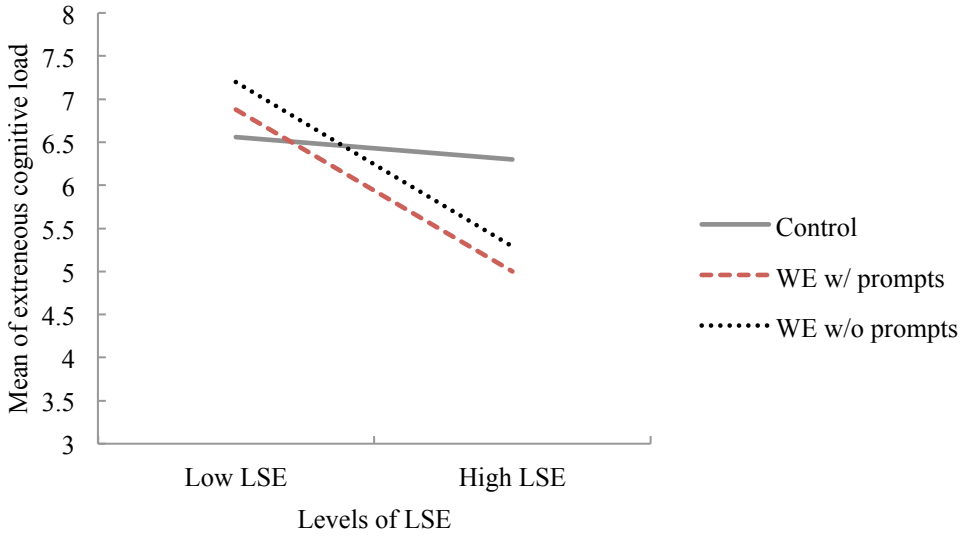

Figure 4-14. The summary of mean values of extraneous cognitive load

Effects on germane cognitive load. A two-way ANOVA was conducted to examine the effect of the treatment and the LSE factor on germane cognitive load. The descriptive statistics of germane cognitive load for each group are shown in Table 4-19.

Table 4-19
The Means and SDs of Germane Cognitive Load Measurements in Each Group

	LSE	*n*	*M*	*SD*
Control	Low LSE	9	5.78	1.30
	High LSE	10	5.80	1.23
WE with prompts	Low LSE	8	7.25	1.49
	High LSE	10	6.70	1.64
WE without prompts	Low LSE	10	7.10	1.10
	High LSE	7	6.43	2.51

As the result of a two-way ANOVA, there was no statistically significant interaction between the treatment factor and the LSE factor, F (2, 48) = 0.259, p = .773, *partial* η^2 = .133. Also there was no main effect of the LSE factor, F (1, 48) = 0.881, p = .353, *partial* η^2 = .018, nor the treatment factor, F (1, 48) = 3.065, p = .056, *partial* η^2 = .113. Figure 4-15 shows the mean of each value.

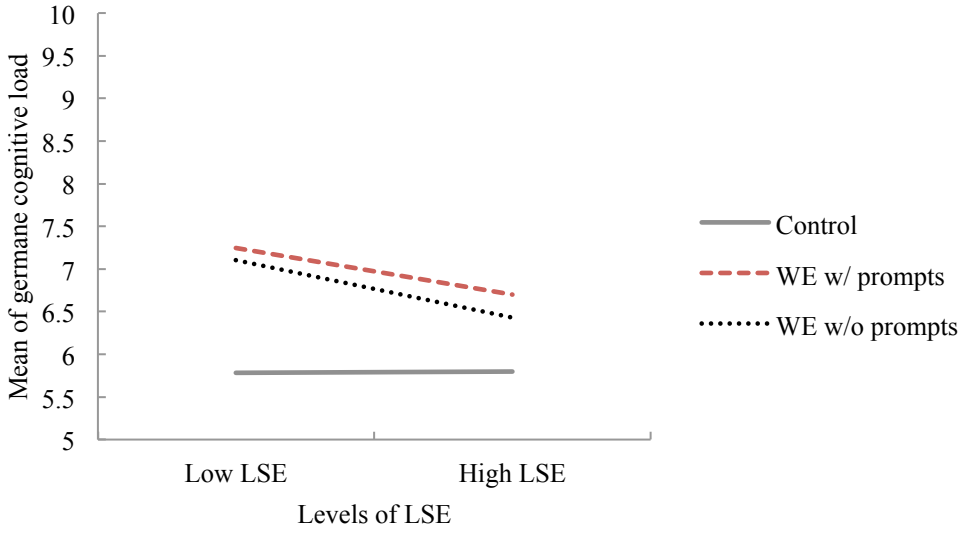

Figure 4-15. The summary of mean values of germane cognitive load

Worked examples and performance in collaborative composition. Regarding Research Question three, one-way ANOVA tests for comparing the mean values of each composition quality criteria and the index of composition complexity for the three treatment groups (Control, WE without prompts, and WE with prompts) were conducted. Descriptive statistics for performance criteria are shown in Table 4-20.

Table 4-20
The Means and SDs of Composition Performance Criteria in Each Group

	Control (*n*=11)		WE without Prompts (*n*=11)		WE with Prompts (*n*=11)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Content	14.82	1.40	17.36	2.50	18.18	2.79
Organization	10.91	2.12	13.09	2.51	14.64	2.16
Vocabulary	10.73	2.20	11.82	2.48	11.64	2.25
Language	10.00	3.44	10.36	3.26	10.27	2.90
Mechanics	2.27	0.47	2.27	0.47	2.64	0.51
GT	48.73	7.08	54.91	7.40	57.36	9.39

As the result of one-way ANOVA tests, the effect of the treatment factor on Content, F (1, 30) = 6.354, p = .005, *partial* η^2 = .298, and Organization, F (1, 30) = 7.497, p = .002, *partial* η^2 = .333, were revealed statistically significant.

On Content, multiple pairwise comparisons using Bonferroni method revealed that the differences in mean values between the Control group and the WE with prompts group and between the Control group and the WE without prompts group were statistically significant, but the differences in mean values between the WE with prompts group and the WE without prompts group were not statistically significant.

On Organization, multiple pairwise comparisons using Bonferroni method revealed that the differences in mean values between the Control group and the WE with prompts group was statistically significant, but the differences in mean values between the Control group and the WE without prompts group, and between the WE with prompts group and the WE without prompts group were not statistically significant.

The rest of criteria, Vocabulary, F (1, 30) = 0.703, p = .503, *partial* η^2 = .045, Language, F (1, 30) = 0.038, p = .962, *partial* η^2 = .003, and Mechanics, F (1, 30) = 2.105, p = .139, *partial* η^2 = .123, did not show statistically significant effects of the treatment factor. Figure 4-16 describes the summary.

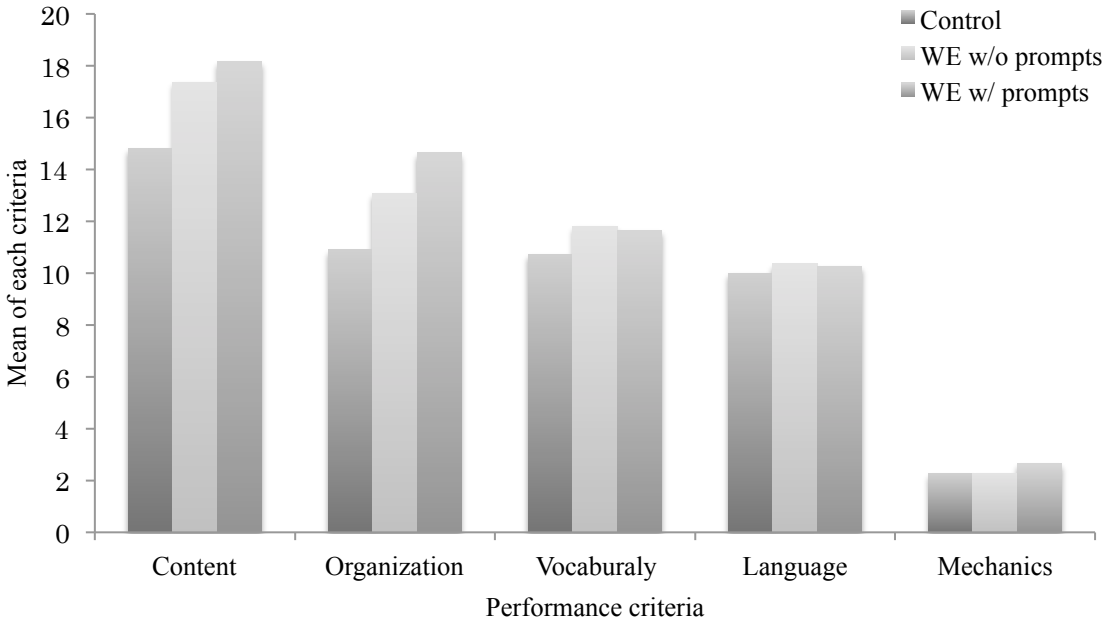

Figure 4-16. The summary of mean composition performance of Study 2

As for the index of complexity of composition, one-way ANOVA did not detect any significant effects of the treatment factor: the number of words, F (1, 30) = 1.598, p = .229, *partial* η^2 = .096, the number of sentences, F (1, 30) = 1.698, p = .200, *partial* η^2 = .102, T-units per sentence, F (1, 30) = 1.113, p = .336, *partial* η^2 = .070. Descriptive statistics of index of complexity are shown in Table 4-21.

Table 4-21
The Comparison of Means by the Treatment Difference on Complexity of Composition

		Control		WE without Prompts		WE with Prompts	
	n	*M*	*SD*	*M*	*SD*	*M*	*SD*
No. of words	11	221.00	94.25	310.00	135.83	301.27	149.53
No. of sentences	11	20.45	6.59	28.00	11.97	26.27	10.82
T-unit per sentence	11	1.16	0.11	1.11	0.08	1.24	0.31

CHAPTER 5 DISCUSSIONS

The two experiments in the present study examined the effects of two types of instructional strategies over four types of stressors uniquely found in online collaboration, three cognitive loads during the collaborative process in relation with language self-efficacy, and collaborative composition performance. Three Research Questions asked were: 1) What type of worked examples would be effective in reducing stressors of EFL learners with two different levels of language self-efficacy in an online collaborative learning context; 2) What type of worked examples would be effective in reducing cognitive loads of EFL learners with two different levels of language self-efficacy in an online collaborative learning context; and 3) What type of worked examples would be effective in contributing to better collaborative performance of EFL learners in an online collaborative learning context?

Study 1 revealed that the WE with prompts was an effective instructional strategy for learners from the low LSE group in an online collaborative learning environment to reduce the influence of stressor; especially Type I stressor (lack of linguistic self-efficacy). Study 1 found the interaction effect between Before-After and treatments on Type I stressor for low LSE group, indicating that EFL learners from the low LSE group had less stress resulted from Type I stressor when using WE with prompts during online collaborative task. Also, Study 1 found the interaction effect between Before-After and treatments on Type IV stressor. Accordingly, it revealed that the group of WE without prompts showed increased level of stress resulted from Type IV stressor in After condition, while it did not find any interaction on Type II and III stressors. As for cognitive loads, Study 1 found an interaction between treatments and LSE on intrinsic cognitive load. The learners from the high LSE group displayed low intrinsic cognitive load when using WE with prompts. However, extraneous and

germane cognitive loads did not find any interaction effects. In addition, composition performance by the alteration of the treatments (WE with prompts or WE without prompts) did not show meaningful difference of between the treatment groups in Study 1.

Study 2 showed contradicting results from Study 1. Namely, there was no substantial difference between the treatment groups in Type I stressor. A meaningful difference was found with the control group. While Study 1 found the low LSE group benefitted from reduction of Type I stressor, Study 2 found the high LSE group to be benefitted. As same as in Study 1, Type II and III stressors were not influenced by treatments or LSE. Type IV stressor did not replicate the same interaction with Study 1. Three cognitive loads did not show any meaningful difference, although intrinsic and extraneous cognitive loads displayed similar influence from the LSE factor. Study 2 detected the differences of composition performance between the control and the treatment groups. Content and Organization were the two criteria found significantly influenced by the provision of worked examples.

Worked Examples and Type I Stressor

Stress mitigation effect of worked example. Study 1 confirmed the effectiveness of WE with prompts as an instructional strategy for mitigating stress by reducing the influence of stressor. That is, it revealed that Type I stressor, one of four types of stressors discussed in Chapter 2, can be reduced using WE with prompts for EFL learners with low LSE during an online collaborative composition task when comparing the Before and After condition. The score of the After condition reduced the amount of Type I stressor significantly from the Before condition.

This can be explained by assuming that the provision of WE with prompts moderates cognitive demand caused by task complexity, thereby mitigating stressor related to that demand. Hence, the finding offers empirical support for the studies by Jung et al. (in press) and Kudo (2013), which recommended the use of worked examples to mitigate Type I stressor during online collaboration. Moreover, it is consistent with the argument made by Niculescu et al. (2009) that cognitive load and stress are related under the construct of cognitive demand; and the argument by Bandura (1995) that self-judgment of one's capability towards the stressors intervenes in the generation of stress reactions.

Consequently, the mitigation of cognitive demand contributes to neutralization of Type I stressor during the course of online collaborative composition. This empirical evidence adds a new characteristic – stress mitigation – for worked examples, in addition to the previously known benefits of reducing extraneous cognitive load, improving learning (e.g., Renkl, 2007; Sweller, 2006; van Gog et al., 2006; van Gog et al., 2011) and fostering acquisition of problem-solving (e.g., Rourke & Sweller, 2009; Rummel & Spada, 2005).

Interaction effect between Before-After and treatments for low LSE learners. The Study 1 further revealed the interaction effect between Before-After and treatments on amount of Type I stressor for the low LSE group. The result showed that the low LSE group benefitted from the WE with prompts in reducing the influence of Type I stressor, while the high LSE group did not. This suggests the levels of LSE regulate the effect of the prompts added to the instructional strategy. In other words, the effect of prompts is determined by the interaction between the learners' expertise level that could be assumed from LSE and the instructional strategy being used. Considering the interaction effect, it implies that prompts has functioned as extra help to allow the

low LSE group to use worked example effectively. The result of Study 1 highlights the advantage of WE with prompts over WE without prompts for the reduction of Type I stressor for low LSE leaners. Learners with low LSE were found to benefit more than their counterparts with high LSE in reducing Type I stressor when using the nine simple self-explaining prompts with the worked example. However, Study 2 did not replicate the results of Study 1. The inconsistency between Study 1 and 2 will be discussed later in the section on *Inconsistent findings and learner characteristics*.

Prompts have been employed as a "strategy activator" (Reigeluth & Stein, 1983) to promote self-regulated learning and consequently enhance learning outcomes. The combination of worked example and prompts was implemented in various domains (e.g., Crippen & Earl, 2007; Hilbert & Renkl, 2009; Hübner et al., 2010). The expectation was that they induce self-explanations, leading to deeper learning. Therefore, the both Study 1 and 2 expected to find WE with prompts inducing similar strategy activation and leading to better performance on collaborative composition. However, the interaction effect suggests that prompts turned out to be a crucial scaffold for learners' linguistic competency.

The two experiments in the present study employed question prompts, which are kind of the textual prompt (Fitzer & Sturmey, 2007). Textual prompts are often used in the context of education; especially online learning condition and themselves represent one of several available presentation formats. Question prompts supply questions to implicitly induce cognitive and metacognitive actions. Through the process of answering these questions, learners notice and realize the intended actions. Question prompts differ from procedural prompts (Scardamalia & Bereiter, 1985; Scardamalia, Bereiter, & Steinbach, 1984), which are characterized by explicit directions or instructions that provoke learner actions.

Further, Ge and Land (2004) argue prompts may even direct students to attend

to that which they would not be able to attend to alone. Prompts use questions to guide students' attention to important information in the learning material. In this manner, question prompts act as scaffolding. Self-explanation prompts are another form of question prompt that elicit learner to self-explain the cognitive process. Self-explanation prompts have been utilized to circumvent a pitfall that learners of low LSE or whom lack pre-existing knowledge often encounter. This is known as the illusion of understanding (cf. Bannert, 2002; Chi et al., 1989), where weaker learners simply skim the surface features of the worked example, leading to superficial learning.

The above framework for conceptualizing the effectiveness of self-explanation prompts is useful for understanding the results of Study 1. The prompts can be considered an extended cognitive tool attached to worked examples and play a pivotal role for learners of low LSE. No such effect was found for the groups of worked examples without prompts. This result is consistent with the argument by Hübner et al. (2010) that both instructional strategy and leaners' level of expertise are influential for the effective use of prompts. Thus, additional prompts acted as a "more able other" (p.15) as argued in Davis (2003). The discussion above implies that the prompts, working as scaffolding, must have enabled the low LSE group to access the benefits of the worked examples during online collaborative composition.

As the two experiments of the present study examined only the effect of question prompts, further studies need to investigate the effect of other types of prompts on the different levels of expertise in EFL online collaborative environments. Other types of prompts might be beneficial for both high and low EFL achievers or they might have different effects depending on the different levels of LSE. Studies in the future must be elaborated to include both EFL competency and LSE.

The effect of LSE. Another aspect of the interaction effect was that the LSE

had an effect on the treatments to influence Type I stressor mitigation. This finding is consistent with the argument made by Pajares (2003) that reading and writing self-efficacy is influential over literal works and the psychological process associated with this. Furthermore, it is in line with Schunk (1989; 2003) claiming the mediating function of self-efficacy that highly efficacious learners participate more readily, work harder, and more tolerant to stressor. Because the experiments are some of few empirically investigate the relationship between stressors for online collaboration and LSE, the finding that self-efficacy is a determining factor for the amount of stressor appears to be unverified. Study 1 revealed that the worked example with prompts is effective only for the low LSE leaners, and worked example alone has limitations. Bandura (1986; 1995; 1997) make a number of arguments including that stress is generated when high 'cognitive demand' is perceived by learners with low self-efficacy, while a high self-efficacy may transform the demands into a challenge. The result of Study 1 provides another layer of empirical evidence to his arguments.

The result from Study 1 is also consistent with findings of CLT studies. From the CLT perspective, Nückles, Hübner, Dümer, and Renkl (2010) discuss the expertise reversal effect (Clark et al., 2006; Reisslein et al., 2006; Renkl, Atkinson, & Große, 2004) that accounts for the difference in reactions to instructional strategy by the levels of LSE. The expertise reversal effect argues for the differences between cognitive expenditure and task design, and explains why a certain instructional strategy is effective for one person, and not effective for the other who has higher knowledge. In the context of Study 1, the expertise reversal effect accounts that the WE with prompts could be effective for the lower LSE who likely obtains low expertise, but the same format of instructional design may be redundant or just extraneous information for the learners of high LSE. It also suggests that effective instructional strategy depends on the interaction of task design and levels of expertise, which is closely related to

self-efficacy. To confirm application of expertise reversal effect on learners' self-efficacy, more empirical evidence is required on task designs and self-efficacy.

Study 2, however, failed to find the replication of Study 1 on Type I stressor. There was an interaction effect between treatments and the LSE, but no difference was found between the two treatment groups. Further explanation is provided for the inconsistent result between Study 1 and 2 in the section on *Inconsistent findings and learner characteristics*.

Worked Examples and Type IV Stressor

In Study 1, the result found Type IV stressor (the stressor stems from the collaborative process itself) increased when using WE without prompts comparing the Before-After condition. This result makes a clear contrast with the finding on Type I stressor that WE with prompts showed reduction in low LSE group. This suggests that WE without prompts experienced the distortion in communication that may have to do with "transactional cost". As previously discussed, the prompts contributed to the effective use of worked example. The increase of Type IV stressor suggests that the absence of the prompts as appropriate scaffoldings might have hampered the smooth communication, leading to a dysfunction in collaboration. It is assumed that the WE without prompts group experienced difficulties utilizing the worked example during collaboration, which in turn, generated higher influence of Type I stressor.

Study 2, however, failed to find the difference of Type IV stressor by the treatments. It found the LSE factor influence the amount of Type IV stressor. The result suggests low LSE groups had difficulties in collaboration. This could be explained by the group difference between Study 1 and 2. Further explanation is provided in the

section on *Inconsistent findings and learner characteristics*.

Worked Examples and Cognitive Loads

Study 1 revealed the interaction effect between treatments and LSE on intrinsic cognitive load. Both treatments and LSE influenced the value of intrinsic cognitive load. Although the difference was not statistically significant, extraneous cognitive load showed an almost identical pattern to that of intrinsic cognitive load. Germane cognitive load was intervened oppositely by LSE, showing that the high LSE group scored higher than the low LSE group.

In Study 2, LSE influenced intrinsic and extraneous cognitive loads. The pattern of values for intrinsic and extraneous cognitive loads was identical to the result of Study 1 indicating the low LSE groups scored higher on both intrinsic and extraneous cognitive loads, except for the control group, which maintained the same values between high and low LSE. There was no meaningful difference between the control group and two treatment groups. Germane cognitive load did not show any meaningful difference.

The interaction effect between treatments and LSE on intrinsic cognitive load fits well with the assumption derived from the past studies. Berthold and Renkl (2009) and Berthold, Röder, Knörzer, Kessler, and Renkl (2011) agreed when the task is complex; the addition of prompts can overload learners' working memory capacity depending on the level of expertise and prior knowledge.

Accordingly, Study 1 found that the high LSE group showed significantly lower intrinsic cognitive load when using WE with prompts than did the low LSE group, while the group of WE without prompts showed no difference between treatments and

LSE. The result confirmed that not only Type I stressor, but also intrinsic cognitive load was influenced by the low LSE when utilizing WE without prompts. This suggests that low LSE group have sensed the increase of task complexity resulting from the addition of prompts, but they have been able to utilize the worked examples with the help of prompts. In contrast, the high LSE group have perceived low intrinsic cognitive load and maintained larger working memory capacity with the help of prompts, despite the increase of task complexity.

The two experiments in the present study also assumed that the low LSE group had smaller working memory and would perceive higher intrinsic cognitive load (Hoffman & Schraw, 2009) when using WE with prompts. The installation of prompts would add one more layer of complexity and processing demand, thus, increasing intrinsic cognitive load. This should be more influential for those who posses smaller working memory capacity. In contrast, the high LSE group would perceive less intrinsic cognitive load when they used WE with prompts because of their larger working memory capacity brought by high LSE and freed up working memory using WE with prompts.

Study 2 revealed a difference by LSE for intrinsic cognitive load. The high LSE group scored lower, and the low LSE group scored higher for intrinsic cognitive load. It failed to find the difference in the alteration of treatments. The addition of prompts did not make a remarkable difference between the treatment groups. The possibility can be that the given task was difficult for the participants of Study 2, so that alternation of treatments did not contribute to reveal the precise difference in the perception of intrinsic cognitive load.

Provided that the worked examples functioned properly, the amount of extraneous cognitive load should have been lowered in the worked example groups.

Both Study 1 and 2 failed to observe such decline of extraneous cognitive load. One possible explanation for this result is the participants' EFL competency. The learners who participated in both studies were EFL false beginners who often find difficult to compose grammatical paragraphs by their own effort. The task difficulty of online collaborative composition was very high for such individuals. Thus, the instructional strategy and additional scaffoldings did not sufficiently support their individual task competency. Therefore, it is assumed that intrinsic cognitive load may be too high; the participants did not differentiate the mental effort brought by the instructional strategies from the task. Further explanation is provided for the inconsistent result between Study 1 and 2 in the section on *Inconsistent findings and learner characteristics*.

Germane cognitive load is associated with processes relating directly to learning, schema construction and automation (van Merriënboer & Ayres, 2005). It is also known as a source of strong motivation (Sweller, 1998). In Study 1, higher degree of germane cognitive load was observed in the high LSE group when compared with the low LSE group. The LSE was shown to be influential on the generation of germane cognitive load. This finding suggests that the cognitive strategies for learning of the high LSE groups were enhanced. However, the alteration of the treatments by addition of prompts failed to show a substantial difference. The hints and questions did not help the WE with prompts group to generate cognitive strategies.

This result seems to contradict with the results on Type I stressor. The Type I stressor was found to be affected by the use of WE with prompts, showing that the WE with prompts group took advantage of prompts. The good use of prompts should lead to the increase of germane cognitive load, on the hind side, it should simultaneously increase intrinsic cognitive load. Unfortunately, Study 2 showed no difference among the three conditions of germane cognitive load. Future studies are called for to explain

why germane cognitive load of the participants' of Study 2 was not influenced by the different treatments.

Worked Examples and Collaborative Composition Performance

Study 2 found that collaborative composition performance, namely the criteria of Content and Organization, were influenced by the difference of treatments, while Study 1 did not find any difference between the treatment groups. The score of Content was significantly higher when comparing with the control group and the two treatment groups, which indicates that worked examples, whether with or without prompts, help improve the quality of content during online collaborative composition. It implies that worked examples are suitable scaffolding strategy for assisting learners with what to write for the collaborative work.

A criterion of Organization was affected by WE with prompts, suggesting prompts were the main source for the improvement of writing organizational structure. This must have been that prompts directed the participants' attention on the structure of the example, and the participants carefully modeled off the worked example. The WE without prompts group did not pay enough attention on the aspect. This implies that addition of prompts can be effective instructional strategy for addressing the leaners' attention for the particular aspects that instructors intend.

The worked example used in the dissertation was a well-written composition that must have helped the participants to borrow their writing Content and Organization, but not influence the Vocabulary, Language use, and Mechanics of the writing. The double-content worked example was beneficial for those low level learners because they can use example as a model and the learners were able to use it for pattern templates

(Renkl, 2013). If the classic worked example, which offers learners with a step-by-step procedure or visual explanation showing experts' procedural tips for problem solving were presented (Renkl, 1997; Sweller et al., 1998), this result would have been different. The findings are in accordance with the effects of prompts that dealt with the structure and contents of the composition, thus it was natural they helped participants focused on Content and Organization of the collaborative composition task, then the treatment groups had significantly higher scores than the control group.

In the end, it should be noted that the use of well-written composition example has hardly been tested under online collaborative composition in EFL context. The recent study that used well-written composition for worked example for literacy work under Korean EFL condition by Kyun et al. (2013) was the one of the similar works. Those positive empirical results will expand the use of worked examples for non-algorithmic fields.

Inconsistent Findings and Learner Characteristics

As reported in the Results chapter, there were several inconsistent findings between Study 1 and 2. The different levels of learner readiness in the task at hand may explain some of these inconsistencies. The participants of Study 2 were freshmen and sophomores who had less collaborative EFL learning opportunities while the participants of Study 1 were junior students with over one more year experience in learning at college. This discrepancy could have affected the quality of the comprehensive language task in both studies in the present study: free composition. The task design was more difficult for the younger participants of Study 2.

As seen in Figure 4-9 in Chapter 4, with regard to reduction of Type I stressor

in Study 2, unlike the result from Study 1, did not reveal a clear difference between two treatment groups – WE with prompts group and WE without prompts group. This inconsistent finding in the effect of prompts on Type I stressor may be accounted by the level of LSE that are thought closely related to the level of expertise. In fact, the participants' LSE was substantially lower in Study 2. The mean score of LSE from Study 1: 1775.00 (the mean of high LSE group = 2538.13 and the mean of low LSE group = 1038.00) was statistically higher than that of Study 2: 1364.00 (the mean of high LSE group = 2101.22 and the mean of low LSE group = 626.19). This significant disparity between the groups may have created such inconsistent results. The extra help provided by the prompts was not enough to scaffold students of low LSE in Study 2.

The fact that the levels of a task domain make a difference in utilizing worked example is agreed with a body of CLT studies. Renkl (2013) and Van Gog and Rummel (2010) attest the lack of prior knowledge may distort the effectiveness of worked examples in English composition. As explained, although all participants in both studies are categorized as false beginners, the disparity of the English grammatical competence between freshmen and juniors are large. Considering the characteristic of false beginners, the level of English grammatical competence must have been substantially lower for the participants of Study 2. Further research is called for to investigate the relationship between the levels of English competency and LSE deeper including the learner variables such as gender, online learning experiences, and learning styles need to be considered in the future studies.

Effective Instructional Strategies for EFL Learners with Low LSE or Novice Learners

The present study hopes that the findings would contribute to provide us with chance to question and reconsider established theories and pedagogies for the online collaboration in the context of EFL. False beginners who lack required skills and/or low self-efficacy cannot produce a functional and rational solution to a given task, especially online collaboration where the learners' self-regulation possesses a great importance. In this case, the provision of worked examples can assist them with showing a concrete written model as an external representation to find a solution or develop a good composition. Furthermore the addition of prompts may support not only effective strategy use for inducing deeper learning, but also reduction of Type I stressor depending on the levels of LSE. It is important to adjust the task design and instructional strategies with learners' personal characteristics (Scheiter, Gerjets, Vollmann, & Catrambone, 2009). Further empirical studies focusing on different applications of worked examples under EFL context adjusting with learners' individual differences should be administered.

The results showed that even EFL learners with low LSE could complete the collaborative composition task. EFL instructors should consider more use of online collaborative composition in EFL classes. Combined with borrowing that has been discussed by EFL experts (Cote, 2006; Hyland, 2003) as traditional instructional strategy of EFL, collaborative composition can be promoted. Although modeling has not been included within the frame of EFL instruction, the results from Study 1 and 2 suggested the possibility of modeling from double-content worked example. Modeling is based on the belief that learning happens when learners borrow and model ideas or behaviors from others. This modeling is likely to increase the probability of successful composition. The research on the modeling of argumentative writing by Braaksma, Rijlaarsdam, and van den Bergh (2002) argued the similarity between learners and the model creates a significant impact on learners' modeling actions and learning results.

This perspective of modeling can be associated with a new worked example

called the double-content examples (Renkl et al., 2009; Schworm & Renkl, 2007; van Gog & Rummel, 2010). The provision of a well-written composition example for written tasks has been recently confirmed effective for supporting deficiency in written competency (Hilbert & Schworm, 2009; Hübner et al., 2010; Kyun et al., 2013; Renkl et al., 2009). The provision of double-content worked examples is an effective instructional strategy for learners of lower self-efficacy; such learners can easily learn the goal state by learning from the worked example as a model. With additional scaffolding strategies, extremely low LSE learners may also be able to benefit from the double-content worked examples. In order to investigate the effectiveness of the double-content worked examples with additional scaffolding strategies, it requires of examinations on strategies effectiveness on extremely low LSE leaners.

It is also important to adjust the levels of complexity of the learning domain so that the learners of variety of LSEs would not encounter trouble in understanding the example. More elaborated studies looking into learners' processes during both collaborative learning are encouraged. Further empirical studies of instructional strategies focusing on different achievement levels, gender, or age group of learners are needed as well, in order to confirm the effect of strategies in EFL online collaborative composition.

CHAPTER 6 CONCLUSION

The present study employed two quasi-experiments designed to examine the effects of two types of instructional strategies (WE with prompts and WE without prompts) on four types of stressors, as well as three kinds of cognitive loads regarding to the two levels of LSE, during online collaborative English composition.

Based on these results it can be concluded that, depending on the levels of LSE, the use of double-content worked examples in combination with self-explanation prompts brings about unique effects during online collaboration of EFL learners. Some inconsistencies arose which necessitate further detailed studies in order to validate this conclusion.

Implications of the Study

Despite some inconsistent findings between Study 1 and Study 2 that have been discussed in Chapter 5, the present study offers useful implications for the design of online collaborative environments, particularly for EFL learners.

- First, it proposes a sound instructional strategy for effective and less stressful online collaboration. Although online collaboration is considered to provide practical and ideal constructivistic learning environments with equal opportunity for various types of learners, it often causes psychological difficulties (Jung et al., 2010; 2012), and creates the issue of "transactional cost" incurred in the collaborative process (Kirschner et al., 2011; Yamane, 1996). The present study suggests that stress will be mitigated with the use of

proper instructional strategies, such as WE with prompts, which has been found to reduce the influence of Type I stressor, especially for the learners of low LSE.

- The present study confirms that EFL false beginners can complete and contribute to the collaborative composition using English with the help of proper instructional strategies. The prompts become an extra help, allowing them to benefit from the use of double-content worked examples in completion of the difficult task by sharing the burden with peers. Online collaboration provides the learners with opportunities for creating a practical context and using foreign language as a means of communication. In light of the findings from the present study, it may be appropriate to reconsider the objectivism-oriented instructional methods.

- The present study informs that any instructional strategy for EFL should consider learners' level of LSE for developing and facilitating online collaborative activities. Because the findings suggest that LSE is influential not only over Type I stressor, but also other stressors and cognitive loads. In online collaboration, where the learners work on their own and collaborate together with their peers, their self-efficacy or confidence affects their participation and performance (Pajares, 1996; 2003).

- The present study suggests that the use of worked examples is effective for EFL collaborative composition administered online. The double-content worked example is a powerful instructional strategy for helping learners to model both structural features and contents from the example, which

consequently mitigates the influence of Type I stressor for the learners of low LSE and maintains low intrinsic cognitive load for the learners of high LSE. Also, the addition of self-explanation prompts can be a support for the use of worked examples when worked examples alone have limitations.

- The present study suggests the importance of developing high self-efficacy, as it is closely related to the reduction of stressors. The instructors' timely encouragements should be incorporated during online collaboration. Positive encouragement for the learners during online collaboration could be a good additional strategy for the promotion of leaners' self-efficacy because one of the four self-efficacy heightening principles (Bandura, 1986) includes social persuasion.

- The present study further offers three suggestions for Japanese EFL teachers.
  1) To use worked examples more in the language classrooms. By utilizing a written model as an external representation, the learners who lack required skills and/or possess low self-efficacy can complete collaborative composition tasks.
  2) To employ more modeling activities using worked examples. Modeling is likely to increase the probability of successful learning when using double-content worked examples.
  3) To enrich traditional EFL borrowing techniques, with the use of double-content worked examples that support deficiency by allowing the learners to readily grasp the ideas from the examples.

- Finally, the present study confirms the importance of scaffolding (Dennen,

2004). The addition of prompts to the worked examples functions as powerful scaffolding for low LSE leaners in Study 1, replicating the effect observed in Hübner et al. (2010). In addition, Study 2 clearly shows the difference in composition performance as well as stress reduction in comparison to the control group. The worked example with prompts has shown strong impact on the low LSE group. The experiments show, whether it is prompted or not, worked examples have contributed to collaborative composition projects for EFL false beginners as scaffolding. It is essential to scaffold learners' lack of confidence with an appropriate support to fill the deficiency of knowledge or skills.

The present study successfully extends the known functions of worked examples to include stressor reduction, targeting one particular type of stressor. The self-explanation prompts have played a crucial role for the low LSE group, allowing them to use worked examples as a model. This particular finding will help EFL instructors to reconsider support for the low LSE group. The results from the present study encourage EFL instructors to use additional scaffolding, so that the learners perceive they have extra help when they engage in online collaborations. As discussed, a variety of prompts, such as procedural prompts that directly instruct procedures for learners to use, should be tested by future studies regardless of additional scaffolding for the lower self-efficacy learners.

Limitations of the Study

While the present study has offered a number of useful guidelines for researchers and

EFL teachers, it has some limitations.

The present study was conducted with intact classes of undergraduates where male students were prevalent. That is, there was a skewed gender balance in the subject groups (about 80% males). Even though Jung et al. (2012) reported that the perceived stressors were not significantly influenced by gender differences, a gender imbalance might have affected the results of Study 1.

There was a possibility that the level of task difficulty perceived by the participants in Study 2 could have been higher than those in Study 1. As explained, the inconsistencies between Study 1 and 2 have been thought to be a result of the discrepancy in English ability between the groups. The participants of Study 2 may have been low in English competency or perhaps the task had been slightly off the difficulty range for the participants. Nevertheless, this might have distorted the findings of the study.

Recommendations for Future Research

Even though the present study has revealed the effect of worked examples with prompts in mitigating a certain type of stressor in online collaboration, there is certainly a need for further research to confirm the results.

- A larger sample size is recommended, in order to understand the effects of instructional strategies. As mentioned in the previous section, the experiments have managed to uncover a statistical significance in the numerical analysis. A larger number of subjects would have helped to increase the validity and credibility of the results.

- Future research should consider experimenting on different institutions and sampling diversified learners' LSE, in order to comprehend the effects of prompts. Testing different prompts or additional scaffolding strategies on different populations may bring about different results. Procedural prompts or directive prompts should be employed for the extremely lower self-efficacy learners. The lower level learners may find them easier to follow.

- The future studies should use grouping based on an absolute criterion for classifying LSE. The inconsistency in the results between Study 1 and 2 suggest the levels of LSE influence reactions to stressors, cognitive loads, and collaborative composition performance due to relative measurement of LSE. More elaborated studies that control the levels of LSE via absolute criterion would provide clear and coherent understanding on the effects of LSE.

- Future studies should include groupings of low and low, high and high, and low and high LSE in order to compare collaborative composition performance. In the present study, performance of collaborative composition has only been evaluated by the treatment differences. Consequently comparison by self-efficacy has not been possible. Since the present study reveals that the levels of LSE are influential over other variables, grouping by the levels of LSE would reveal this phenomenon in more detail.

- Future studies should attempt to administer a longitudinal experiment in order to observe the change of stressors over time, under a more rigidly controlled experimental condition. The present study has used a quasi-experiment

method where college students engage in online collaboration for only a short period of time (four weeks). Further, the longer-period study should focus not only on numerical data, but also on the collaborative activities of each participant and in-depth interviews regarding the experience of the collaboration. All of this will be more achievable under experimental conditions.

- Future studies need to include a more gender-balanced design. Although a previous study (Jung et al., 2011) argues gender balance does not affect perception of stressors, much EFL literature indicates gender plays an important role in written communication (Kobayashi, 2002). In addition to gender, online learning experiences and learning styles as individual differences are recommended for further research.

- Future studies are recommended to use different tasks (such as writing a research paper, debating arguments, or drafting legal documents). Task difficulty must be controlled in order to avoid the risk of heavy stress induction. When learners have trouble comprehending the exemplifying domain (the contents) of double-content worked examples, they encounter another layer of difficulty. Controlling task difficulty would provide deeper and more precise confirmation of the effects of instructional strategies.

REFERENCES

Akgun, S., & Ciarrochi, J. (2003). Learned resourcefulness moderates the relationship between academic stress and academic performance. *Journal of Educational Psychology, 23*(3), 287-294.

Alavi, M. (1994). Computer-mediated collaborative learning: An empirical evaluation. *MIS Quarterly, 18*(2), 59-174.

Al-Fudail, M., & Mellar, H. (2008). Investigating teacher stress when using technology. *Journal of computer and science, 51,* 1103-1110.

Allan, J., & Lawless, N. (2003). Stress caused by on-line collaboration in e-learning: A developing model. *Education and Training, 45*(8-9), 564-572.

Ananiadou, K., & Claro, M. (2009). 21st century skills and competences for new millennium learners in OECD countries. *OECD Education Working Papers*, No. 41. Paris: OECD Publishing. doi: 10.1787/218525261154

Arai, K. (2004). *"Naniga gaikokugo gakushusha no yaruki wo ushinawasewruka?: Douki gentai no genin to sorenitaisuru gakushuusha no hannou nikansuru shituteki chousa".* [What demotivates foreign language leaners?: A qualitative survey on the cause of demotivation and the leaners' reactions]. The bulletin Toyogakuen University, *12*, 39-47.

Arnold, N., & Ducate, L. (2006). Future foreign language teachers' social and cognitive collaboration in an online environment. *Language Learning & Technology, 10*(1), 42-66.

Asher, J.J. (1972). Children's first language as a model for second language learning. *The Modern Language Journal, 56*(3), 3-17.

Atkinson, R.K., Derry, S.J., Renkl, A., & Wortham, D.W. (2000). Learning from examples: instructional principles from the worked examples research. *Review of Educational Research, 70,* 181-214.

Atkinson, R.K., Renkl, A., & Merrill, M.M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology, 95*(4), 774-783.

Ayres, P., & Paas, F. (2009). Interdisciplinary perspectives inspiring a new generation of cognitive load research. *Educational Psychology Review, 21,* 1-9

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147.

Bandura, A. (1986). *Social foundations of thought and action: A social- cognitive theory.* Engelwood Cliffs, NJ: Prentice Hall.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational psychologist*, *28*(2), 117-148.

Bandura, A. (Ed.). (1995). *Self-efficacy in changing societies*. Cambridge University Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: New York: W. H. Freeman and Company.

Bandura, A., & Cervone, D. (1986). Differential engagement of self-reactive influences in cognitive motivation. *Organizational Behavior and Human Decision Processes, 38,* 92-113.

Barkley, E.F., Cross, K.P., & Major, C.H. (2005). *Collaborative learning techniques: A handbook for college faculty.* San Francisco, CA: Jossey-Bass.

Bannert, M. (2002). Managing cognitive load – recent trends in cognitive load theory. *Learning and Instruction, 12*, 139-146.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*, 276-292.

Beckmann, J. F. (2010). Taming a beast of burden – On some issues with the conceptualization and operationalisation of cognitive load. *Learning and Instruction, 20,* 250-264.

Bernard, R.M., Rojo de Rubacava, B., & St-Pierre, D. (2000). Collaborative online distance learning: Issues for future practice and research. *Journal of Distance Education*, *21*(2), 260-277.

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology*, *101*(1), 70-87.

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*, 564-577.

Berthold, K., Röder, H., Knörzer, D., Kessler, W., & Renkl, A. (2011). The double-edged effects of explanation prompts. *Computers in Human Behavior, 27,* 69-75.

Borg, G., Bratfisch, O., & Dornic, S. (1971). On the problem of perceived difficulty. *Scandinavian Journal of Psychology, 12*, 249-260.

Braaksma, M.A.H., Rijlaarsdam, G., & van den Bergh, H. (2002). Observational learning and the effects on model-observer similarity. *Journal of Educational Psychology, 94,* 405-415.

Brindley, J., Walti, C., & Blaschke, L. (2009). Creating effective collaborative learning groups in an online environment. *The International Review of Research In Open And Distance Learning, 10*(3), Article 10.3.11. Retrieved Nov. 22, 2011. from http://www.irrodl.org/index.php/irrodl/article/view/675/1271

Brünken, R., Plass, J.L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53-61.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8,* 293-332.

Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145-182.

Chi, M.T.H., de Leeuw, N., Chiu, M.H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477.

Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior, 25,* 315-324.

Clark, T. (2003). Disadvantages of collaborative online discussion and the advantages of sociability, fun and cliques for online learning. *Conferences in Research and Practice in Information Technology, Australia,* 23-25.

Clark, R., Nguyen, F., & Sweller, J. (2006). *Efficiency in Learning: evidence-based guidelines to manage cognitive load.* Pfieffer: San Francisco, CA.

Cohen, S., & Gibson, C. (2003). *Mutual understanding, integration and trust: Creating conditions for virtual team effectiveness.* Unpublished Paper, Marshall School of Business, University of Southern California.

Cole, M. (2009). Using Wiki technology to support student engagement: Lessons from the trenches. *Computers & Education, 52,* 141-146.

Cote, R. (2006). *Peer collaboration in the ESL writing classroom: A literature synthesis.* Retrieved Nov. 14, 2011. from http://www.u.arizona.edu/~rcote/SLAT596O/Term%20Paper.pdf.

Crippen, K. J., & Earl, B. L. (2004). Considering the efficacy of web-based worked examples in introductory chemistry. *Journal of Computers in Mathematics and Science Teaching, 23*(2), 151-167.

Crippen, K.J., & Earl, B.L. (2007). The impact of web-based worked examples and self-explanation on performance, problem solving, and self-efficacy. *Computers & Education, 49*, 809-821.

Davis, E. (2003). Prompting middle school science students for productive reflection: generic and directed prompts, *Journal of the Learning Sciences, 12,* 91-142.

Davis, E.A., & Linn, M.C. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education, 22,* 819-837.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science, 38,* 105-134.

Dennen, V.P. (2004). Cognitive apprenticeship in educational practice: Research on scaffolding, modeling, mentoring, and coaching as instructional strategies. In Jonassen D.H. (Ed). *Handbook of Research on Educational Communications and Technology.* 2nd ed. pp. 813-828 Mahwah NJ.: Lawrence Erlbaum Associates.

Diao, Y., & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations, *Learning and Instruction, 17*, 78-88.

Dillenbourg, P., & Schneider, D. (1995). *Collaborative learning and the Internet.* Retrieved Nov. 26, 2011. from http://tecfa.unige.ch/tecfa/research/CMC/colla/iccai95_1.html

Dörnyei, Z. (1998). *Demotivation in foreign language learning.* Paper presented at the TESOL '98 Congress, Seattle, WA, March.

Ehrmana, M.E., Leaverb, B.L., & Oxford, R.L. (2003). A brief overview of individual differences in second language learning. *System, 31,* 313-330.

Fitzer, A., & Sturmey, P. (2007). *Autism spectrum disorders: Applied behavior analysis, evidence, and practice.* In W.H. Ahearn, W.V. Dube, R. MacDonald, & R.B. Graff. (Eds.), Behavior analytic teaching procedures: Basic principles, empirically derived practices (pp. 31-72). Austin, TX: PRO-ED.

Fjermestad, J. (2004). An analysis of communication mode in group support systems research. *Decision Support Systems, 37*(2), 239-263.

Folkman, S., & Moskovitz, J.T. (2000). Stress, positive emotion, and coping. *Current Directions in Psychological Science, 9*(4), 115-118.

Gaillard, A.W.K. (1993). Comparing the concepts of mental load and stress. *Ergonomics, 36*(9), 991-1005.

Garrison, D.R. (2006). Online collaboration principles. *Journal of Asynchronous Learning Networks, 10*(1), 25-34.

Ge, X., & Land, S.M. (2004). A conceptual framework for scaffolding ill-structured problem-solving processes using question prompts and peer interactions. *Educational Technology Research and Development, 52*(2), 5-22.

Gerlach, J.M. (1994). Is this collaboration. In K. Bosworth & S.J. Hamilton (Eds.) *Collaborative Learning: Underlying Processes and Effective Techniques. New Directions for Teaching and Learning, No. 59.* San Francisco: Jossey-Bass Publishers.

Gillani, B.B. (2003). *Learning theories and the design of e-learning environments.* New York: University Press of America.

Guldberg, K., & Pilkington, R. (2006). A community of practice approach to the development of non-traditional learners through networked learning. *Journal of Computer Assisted Learning, 22,* 159-171.

Gunawardena, C.N. (1995). Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications, 1(2/3),* 147-166.

Helgesen, M. (1991). Teaching in Japan. *TESOL Journal, 1*(1), 8.

Hilbert, T.S., Renkl, A., Schworm, S., Kessler, S., & Reiss, K. (2008). Learning to teach with worked-out examples: A computer-based learning environment for teachers. *Journal of Computer Assisted Learning, 24,* 316-332. doi:10.1111/ j.1365-2729.2007.00266.x.

Hilbert, T.S., & Renkl, A. (2009). Learning how to use a computer-based concept-mapping tool: Self-explaining examples helps. *Computers in Human Behavior, 25,* 267-274.

Hoffman, B., & Schraw, G. (2009). The influence of self-efficacy and working memory capacity on problem-solving efficiency. *Learning and Individual Differences, 19(1),* 91-100.

Hogue, A. (1996). *First Steps in Academic Writing.* Addison-Wesley Longman Publishing Company. White Plains, NY.

Hsiao, J.D.L. (2005). *CSCL theories.* Retrieved Dec. 6, 2011. From http://www.edb.utexas.edu/cscl/Dhsiao/theories.html

Hunt, K. (1996). Recent measures in syntactic development. *Elementary English, 43,* 139-732.

Hübner, S., Nückles, M., & Renkl, A. (2010). Writing learning journals: Instructional support to overcome learning-strategy deficits. *Learning and Instruction, 20*(1), 18-29.

Hyland, K. (2003). *Second language writing.* Cambridge: Cambridge University Press.

Ingram, A.L., & Hathorn, L.G. (2003). Designing your Web site for instructional effectiveness and completeness: First step. *Tech Trends, 47*(2), 50-56.

Jacobs, H., Hartfiel, V. Hughey, J., & Wormuth, D. (1981). Testing ESL Composition: A. Jacobs, H (Ed). *Testing ESL composition: A practical approach.* Rowley, MA: Newbury House.

Johnson, D.W., & Johnson, R.T. (1999). *Learning together and alone: cooperative, competitive, and individualistic learning (5th ed.).* Boston: Allyn & Bacon.

Johnson, D.W., & Johnson, R.T. (2004). Cooperation and the use of technology. In D.H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 785-811). Mahwah, NJ: Erlbaum.

Johnson, D.W., Johnson, R., & Smith, K. (1998). *Active learning: Cooperation in the college classroom.* Edina, MN: Interaction Book Company.

Johnson, M. (2010). An Exploratory Study of Japanese Engineering Students' EFL Learning Motivation. *The Proceedings of HICE*, 2586-2601.

Jung, I.S., Kudo, M., & Choi, S.K. (2011). Instructional design strategies for stress-reduced online collaboration in Asia's high context culture. *Proceedings of the 25th annual AAOU conference.*

Jung, I.S., Kudo, M., & Choi, S.K. (2012). Stress in Japanese learners engaged in online collaborative learning in English. *British Journal of Educational Technology. 43*(6). 1016-1029. doi:10.1111/j.1467-8535.2011.01271.x

Jung, I.S., Kudo, M., & Choi, S.K. (in press). Towards effective and less stressful online collaborative learning: Strategies to promote engagement while minimizing unnecessary cognitive load and stress. *The International Handbook of E-learning.* Athabasca University Press.

Kalyuga, S. (2007). Expertise Reversal Effect and Its Implications for Learner-Tailored Instruction. *Educational Psychological Review, 19*, 509-539.

Kalyuga, S. (2008). Schema acquisition and sources of cognitive load. In J. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive Load: Theory & Application.* (pp. 48-64). New York, NY: Cambridge University Press.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40,* 1-17.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). Expertise reversal effect. *Educational Psychologist, 38*(1), 23-31.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 351-371.

Kameyama, T. (2007). Kousen-sei notameno hinshirikai to koubunnryoku tono kannenn ni chakumokushita eigokoubun gakushuyou online kyouzai. [Online instructional materials for learning English structures aiming at comprehension over parts of speech and structures for the students of technical colleges.] *The Bulletin of Technical College Education 2007,* 251-254.

Kanatani, K. (2004). *Koukou eigowo kaeru wayaku sakiwatashi jugyono kokoromi. [A trial of early handout of Japanese translation to change English education at high schools].* Sanseido: Tokyo.

Koedinger, K.R., & Aleven V. (2007). Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review, 19*(3)*,* 239-264.

Kobayashi, Y. (2002). The role of gender in foreign language learning attitudes: Japanese female students' attitudes English learning. *Gender and Education*, *14*, 181-197.

Krause, U. M., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction*, *19*(2), 158-170.

Kudo, M. (2007). Searching for an alternative instructional strategy for remedial EFL online learning: based on Cognitive Load Theory. *The Proceedings of E-learn 2007 conference,* 1663-1668.

Kudo, M. (2013). Cognitive load theory (CLT) as a conceptual and practical means for improving English as a foreign language education in Japan. *Educational Studies, 55,* 55-62.

Kudo, M., Choi, S.K., & Jung, I.S. (2010). Toward a model of stress in second language-based online collaborative learning. *Proceedings of the 26th Annual Conference of JSET,* 1001-1002.

Kudo, M., Ohbuchi, K., & Matsui, M. (2003). The collaboration between Information Technology and English – an interdisciplinary approach for developing CALL application. *The Bulletin of SSTA, 12,* 57-65.

Kurokami, H., Horita, T., Yamauchi, Y., Inagaki, T., & Minowa, T. (2001). A small history of a communication board for collaborative among distant elementary classes. In C. Montgomerie, & J. Viteli (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2001* (pp. 1035-1040). Chesapeake, VA: AACE.

Kimura, Y., Nakata, Y., & Okumura, T. (2001). Language learning motivation of EFL learners in Japan*:* A cross-sectional analysis of various learning milieus. *Japan Association for Language Teaching Journal, 23*(1), 47-68.

Kirschner, P.A. (2001). Cognitive load theory: implications of cognitive load theory on the design of learning. *Learning and Instruction, 10*(Suppl. 1)*,* 1-9.

Kirschner, P.A. (2002). Using integrated electronic environments for collaborative teaching/learning. *Learning and Instruction, 12,* 1-10.

Kirschner, F., Paas, F., & Kirschner, P.A. (2009a). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior, 25*, 306-314.

Kirschner, F., Paas, F., & Kirschner, P.A. (2009b). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review, 21*, 31-42.

Kirschner, F., Paas, F., & Kirschner, P.A. (2011). Superiority of collaborative learning with complex tasks: A research note on an alternative affective explanation. *Computers in Human Behavior, 27,* 53-57.

Kirschner, F., Paas, F., Kirschner, P.A., & Janssen, J. (2011). Differential effects of problem-solving demands on individual and collaborative learning outcomes. *Learning and Instruction, 21*, 587-599.

Kyun, S., Kalyuga, S., & Sweller, J. (2013). The Effect of Worked Examples When Learning to Write Essays in English Literature, *The Journal of Experimental Education, 81(3),* 385-408, DOI: 10.1080/00220973.2012.727884

Lazarus, R.S. (1990). Theory-based stress measurement. *Psychological Inquiry, 1*(1), 3-13.

Lazarus, R.S., & Folkman, S. (1984). *Stress, Appraisal, and Coping.* New York, NY: Springer.

Lawless, N., & Allan, J. (2004). Understanding and reducing stress in collaborative e-Learning. *Electronic Journal of e-Learning, 2*(1), 121-127.

Levy, M. (2009). Web 2.0 implications on knowledge management. J*ournal of Knowledge Management, 13,* 120-134.

Linnenbrink, E.A., & Pintrich, P.R. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading and Writing Quarterly, 19*, 119-137.

Lipponen, L. (1999). The challenges for computer supported collaborative learning in elementary and secondary level: Finnish perspectives. In C.Hoadley (Ed.), *Proceedings of The Third International Conference on Computer Support for Collaborative Learning* (pp. 368-375). NJ: Lawrence Erlbaum Associates.

McConnell, D. (2000). *Implementing computer supported cooperative learning.* London: Kogan Page.

Mäkitalo, K., Weinberger, A., Häkkinen, P., Järvelä, S., & Fischer, F. (2005). Epistemic cooperation scripts in online learning environments: fostering learning by reducing uncertainty in discourse? *Computers in Human Behavior, 21,* 603-622.

Mayer, R.E. (2001). *Multimedia learning.* New York: Cambridge University Press.

Mayer, R.E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction, 12*(1), 107-119.

McCarthey, S. J., & McMahon, S. (1992). From convention to invention: Three approaches to peer interactions during writing. In R.Hertz-Lazarowitz, & N. Miller (Eds.), *Interaction in cooperative groups.* The theoretical anatomy of group learning (pp.17-35). Cambridge: Cambridge University Press.

Miller, G.A. (1956). The magical number seven, plus minus two: Some limits on our capacity for processing information. *Psychological Review, 14,* 263-264.

Mills, N. (2009). A Guide du Routard simulation: Increasing self-efficacy in the standards through project-based learning. *Foreign Language Annals, 42*(4), 607-639.

Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: self-efficacy, anxiety and their relation to reading and listening proficiency. *Foreign Language Annals, 39*(2), 273-292.

Mills, N., Pajares, F., & Herron, C. (2007). Self-efficacy of college intermediate French students: relation to achievement and motivation. *Language Learning, 57*(3), 417-442.

Morgan, R.L., Whorton, J.E., & Gunsalus, C. (2000). A comparison of short-term and long-term retention: Lecture combined with discussion versus cooperative learning. *Journal of Instructional Psychology, 27,* 53-58.

Morley, J. (1990). Trends and developments in listening comprehension: Theory and practice. *Georgetown University Round Table on Languages and Linguistics.* 317-337.

Moreno, R. (2004). Decreasing Cognitive Load for Novice Students: Effects of Explanatory versus Corrective Feedback in Discovery-Based Multimedia. *Instructional Science, 32,* 99-113.

Mulligan, C., & Garofalo, R. (2011). A collaborative writing approach: Methodology and student assessment. *The Language Teacher, 35*(3), 5-10.

Nakamura, A. (2005). *A proposal for remedial education in collage. [Daigaku niokeru remedial kyouiku heno teigen]*. Daigakukyoiku Shuppan: Okayama City.

Niculescu, A.I., Cao, Y., & Nijholt, A. (2009). Stress and cognitive load in multimodal conversational interactions. *Proceedings of 13th International Conference on Human-Computer Interaction,* In C. Stephanides (Eds), Springer Verlag, Heidelberg , 891-895. ISBN 978-3-642-02884-7.

Nückles, M., Hübner, S., Dümer, S., & Renkl, A. (2010). Expertise reversal effects in writing-to-learn. *Instructional Science*, *38*(3), 237-258.

Palloff, R.M., & Pratt, K. (2005). *Collaborating online – Learning Together in Community.* Josset-Bass, San Francisco, CA.

Palmer, S., Cooper, C., & Thomas, K. (2003). *Creating a Balance: Managing Stress.* London: British Library.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent Developments. *Educational Psychologist, 38(1),* 1-4.

Paas, F., & van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*(1), 122-133.

Paas, F., Van Merriënboer, J., & Adam, J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79,* 419-430.

Paas, F., Tuovinen, J., Tabbers, H., & Van Gerven, P. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-71.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading &Writing Quarterly*, *19*(2), 139-158.

Pajares, F., & Miller, D.M. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology, 86,* 193-203.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction, 12(1),* 61-86.

Retnowati, E., Ayres, P., & Sweller, J. (2010). Worked example effects in individual and group work settings. *Educational Psychology, 30*(3), 349-367.

Reigeluth, C.M., & Stein, F.S. (1983). The elaboration theory of instruction. In C.M. Reigeluth (Ed.), *Instructional-design theories and models: an overview of their current status* (pp. 335-382). Hillsdale, NJ: Erlbaum.

Reisslein, J., Atkinson, R.K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction*, *16*(2), 92-103.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21*, 1-29.

Renkl, A. (2002). Worked-out examples: instructional explanations support learning by self-explanations. *Learning and Instruction, 12*, 529-556.

Renkl, A. (2005). The worked-out-example principle in multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp.229-246). Cambridge: Cambridge University Press.

Renkl, A. (2013). Toward an Instructionally Oriented Theory of Example-Based Learning. *Cognitive Science, 37,* 1-37.

Renkl, A., Atkinson, R.K., & Große, C.S. (2004). How fading worked solution steps works: A cognitive load perspective. *Instructional Science, 32*, 529-556.

Renkl, A., & Atkinson, R. K. (2007). *An example order for cognitive skill acquisition.* In F.E. Ritter, J. Nerb, E. Lehtinen, & T.M. O'Shea (Eds.), In order to learn: how the sequence of topics influences learning (pp. 95-105). New York: Oxford University Press.

Renkl, A., Hilbert, T., & Schworm, S. (2009). Example-based learning in heuristic domains: A cognitive load theory account. *Educational Psychology Review, 21,* 67-78.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23*, 90-108.

Rourke, A., & Sweller, J. (2009). The worked-example effect using ill-defined problems: Learning to recognize designers' styles. *Learning and Instruction, 19,* 185-199.

Rummel, N., & Spada, H. (2005). Learning to collaborate: an instructional approach to promoting collaborative problem-solving in computer-mediated settings. *Journal of the Learning Sciences, 14,* 201-241.

Rummel, N., Spada, H., & Hauser, S. (2006). Learning to collaborate in a computer-mediated setting: observing a model beats learning from being scripted. In S. Barab, K. Hay, & D. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences. (pp. 634–640)*. Mahwah: Erlbaum.

Salmon, G.K., Allan, J.S., & Giles, K. (2000). Training for online working. In K. Daniel, D. Lamond, & P. Standen (Eds.), *Managing telework.* London: Thompson.

Sawyer, R.K. (2006). Introduction: the new science of learning. In R.K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 1-16). New York, NY: Cambridge.

Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. *Cognitive Science, 8,* 173–190.

Scardamalia, M., & Bereiter, C., 1987. Knowledge telling and knowledge transforming in written composition. In Rosenberg (Ed.). Reading, Writing and Language Learning. *Advanced in Applied Psycholinguistics, 2*, (pp. 142-175). Cambridge: Cambridge University Press.

Scheiter, K., Gerjets, P., Vollmann, B., & Catrambone, R. (2009). The impact of learner characteristics on information utilization strategies, cognitive load experienced, and performance in hypermedia learning. *Learning and Instruction, 19*, 387-401.

Schunk, D.H. (1989). Self-efficacy and cognitive achievement: Implications for students with learning problems. *Journal of Learning Disabilities, 22*, 14-22.

Schunk, D.H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting and self-evaluation. *Reading & Writing Quarterly, 19,* 159-172.

Schworm, S., & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology, 99,* 285-296.

Shaw, E.J. (2007). *The reading and writing self-efficacy beliefs of students with discrepant reading and writing performance.* Unpublished Doctoral Dissertation, Fordham University.

Shell, D.F., Colvin, C., & Bruning, R.H. (1995). Self-efficacy, attribution, and outcome expectancy mechanisms in reading and writing achievement: Grade- level and achievement-level differences. *Journal of Educational Psychology, 87,* 386-398.

Shepherd, J. (2008, January 15). Desperately seeking students. The Guardian. Retrieved Nov. 26, 2011. from http://www.theguardian.com/education/2008/jan/15/internationaleducationnews.highereducation/

Simpson, O. (2000). *Supporting students in open and distance learning.* London: Kogan Page.

Spanjers, I., van Gog, T., & van Merriënboer, J. (2012). Segmentation of Worked Examples: Effects on Cognitive Load and Learning. *Applied Cognitive Psychology, 26,* 352-358.

Spanjers, I., Wouters, P., van Gog, T., & van Merriënboer, J. (2011). An expertise reversal effect of segmentation in learning from animated worked-out examples. *Computers in Human Behavior, 27,* 46-52.

Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing, 14,* 153-173.

Stark, R., Kopp, V., & Fischer, M.R. (2011). Case-based learning with worked examples in complex domains: Two experimental studies in undergraduate medical education. *Learning and Instruction, 21*, 22-33.

Strijbos, J.W., Martens, R.L., & Jochems, W.M.G. (2004). Designing for interaction: Six steps to designing computer-supported group-based learning. *Computers & Education, 42,* 403-424.

Suzuki, K. (2000). *The dictionary of Educational Technology. [Kyoiku kougaku jiten]*. Gihyousha: Tokyo.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* 257-285.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology, 81*(4), 457-466.

Sweller, J. (1999). *Instructional design in technical areas.* Camberwell, Australia: ACER Press.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The Psychology of Learning and Motivation, 43,* pp. 215-266. Academic Press: San Diego.

Sweller, J. (2006). The worked example effect and human cognition. *Learning and Instruction, 16*(2) 165-169.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory.* Springer: NY, NY.

Sweller, J., & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2,* 59-89.

Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology, 4,* 434-458.

Sweller, J., & Levine, M. (1982). Effects of goal specificity on means-ends analysis and learning. Journal of Experimental Psychology. *Learning, Memory, and Cognition, 8,* 463-474.

Sweller, J., Van Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Takefuta, Y., & Suiko, M. (2005). *Korekarano Daigaku Eigokyouiku* [The Prospective English education in Japan]. Iwanami Pub: Tokyo.

Tarmizi, R.A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*(4), 424-436.

Thomas, J. (2006). Studying Le Ballon Rouge with false beginners. *Academic Exchange Quarterly, 10*. Retrieved Oct. 13, 2013 from http://findarticles.com/p/articles/mi_hb3325/ is_1_10/?tag=content;coll

Thomée, S., Eklöf, M., Gustafsson, E., Nilsson, R., & Hagberg, M. (2007). Prevalence of perceived stress, symptoms of depression and sleep disturbances in relation to information and communication technology (ICT) use among young adults – an explorative prospective study. *Computers in Human Behavior, 23,* 1300-1321.

Tobias, S. (1987). Mandatory text review and interaction with student characteristics. *Journal of Educational Psychology, 79*, 154-161.

Trafton, J.G., & Reiser, B.J. (1993). The contributions of studying examples and solving problems to skill acquisition. In M. Poison (Ed.), *Proceedings of the Fifteenth Annual Conference ~ the Cognitive Science Society* (pp. 1017-1022). Hillsdale, NJ: Erlbaum.

University of California, Barkley. (2013). Graduate Students Instructor Teaching & Resource Center. Retrieved Oct. 13, 2013, from http://gsi.berkeley.edu/teachingguide/theories/overview.html

Van Bruggen, J., Kirschner, P.A., & Jochems, W. (2002). External representation of argumentation in CSCL and the management of cognitive load. *Learning and Instruction, 12,* 121-138.

van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22,* 155-174.

Van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example-problem, and problem-example pairs on novices' learning. *Contemporary Educational Psychology, 36*, 212-218.

van Gog, T., Paas, F., & van Merriënboer, J.J.G. (2006). Effects of process-oriented worked examples on troubleshooting transfer performance. *Learning and Instruction, 16,* 154-164.

van Merriënboer, J. (1997). *Training Complex Cognitive Skills.* Educational Technology Publications, Englewood Cliffs: New Jersey.

Van Merriënboer, J., & Ayres, P. (2005). Research on cognitive load theory and its design imprecations for e-learning. *Educational Technology Research and Development, 53*(3), 5-13.

Van Merrienboer, J.J.G., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology, 20,* 343-352.

Van Merrienboer, J.J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational psychology review*, *17*(2), 147-177.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological process.* Cambridge: Harvard University Press.

Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction, 7,* 1-39.

Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety, *System*, doi:10.1016/ j.system.2011.10.017

Yamane, D. (1996). Collaboration and its discontents: steps toward overcoming barriers to successful group projects. *Teaching Sociology, 24*, 378-383.

Yamanishi, H. (2004). "Koukousei no Jiyueisakubun ha donoyouni hyouka sareteirunoka – bunsekiteki shakudo to sougouteki hyouka shakudo no hikaku wo toshiteno kento" [How are High School Students' Free Compositions Evaluated by teachers and Teacher Candidates?: A Comparative Analysis between Analytic and Holistic Rating Scales]. *JALT Journal, 26*(2), 189-206.

Yoshida, K. (2004). *"Eigoga tsukaeru nihonjinn" no ikuseino tameno eigokyouin kenshu guidebook.* [A guidebook for teacher trainings to raise "Japanese who can use English"]. Kairyudo: Tokyo.

Zajacova, A., Lynch, S.M., & Espenshade, T.J. (2005). Self-efficacy, stress, and academic success in college. *Research in Higher Education, 46(6),* 677-705.

Zang, K., & Peck, K.L. (2003). The effects of peer-controlled or moderated online collaboration on group problem solving and related attitudes. *Canadian Journal of Learning and Technology, 29*(3). Retrieved Oct. 13, 2013 from http://www.cjlt.ca/index.php/cjlt/article/view/91/85

APPENDICES

Appendix A: Stressor scale

Translated from Jung, I.S., Kudo, M., & Choi, S.K. (2012). Stress in Japanese learners engaged in online collaborative learning in English. *British Journal of Educational Technology*. *43*(6). 1016-1029.

1. 学科: _____
2. 性別: __ 女性 __ 男性
3. 氏名: _____
4. 学生番号: _____
5. オンライン学習（例: Moodle を使ったディスカッション、またはオンライン協働課題の経験など）に関して、あなたの経験を評価して下さい。
   - □ 全くなし
   - □ 少しある
   - □ 多少ある
   - □ 多い
   - □ かなり多い
6. あなたは、他の学生との協働課題に取り組むことを楽しみますか？
   - □ 全く楽しくない
   - □ 少し楽しい
   - □ ある程度楽しい
   - □ ほとんどの場合楽しい
   - □ とても楽しい
7. どのようなモバイル機器を使っていますか？（複数回答可）
   - □ スマートフォン (例: iPhone, Adroid phone, Gallaxy)
   - □ 普通の携帯 (スマートフォン以外)
   - □ タブレット PC (例: iPad, Gallaxy Tab)
   - □ E-book リーダー (例: Kindle, Sony's Librie)
   - □ ラップトップコンピュータ
   - □ 携帯ゲーム機器 (例 Nintendo DS, PlayStation Portable)
   - □ その他 (具体名:)
8. 今回の課題でどれくらい授業外での勉強時間が増えましたか？
   - □ 1 時間程度
   - □ 2 時間程度
   - □ 3 時間程度
   - □ 4 時間程度
   - □ 5 時間以上

なぜオンライン協働環境でストレスを感じるのでしょう？

・この調査は、英語を使ってオンライン協働学習を行う際に、どのようなストレス要因があるかを明確にするためのものです。この調査によって確認されたストレス要因をもとにして、できるだけストレスの少ない英語オンライン協働学習を計画することができるようになります。

・この調査への参加は任意であり、いつ回答をやめてかまいません。あなたの回答は、他の回答と一緒となり、すべての個人的データは外部に公表されることはありません。この調査に参加することであなたへの不利益はまったくありません。あなたの協力が、ストレスの少ないオンライン協働学習を作り上げるために重要な貢献となります。

・この調査は、すべて回答すると 10 分程度で終了します。回答の完了と提出をもって、この研究への参加の意志表明となります。

ストレス要因
Moodle などを使ってグループ課題に取り組むとします。このようなオンライン協働学習環境下で、何がストレスをもたらすと考えますか？以下の項目に関して、あなたにとってどのくらいストレスがありますか。1~5 からあてはまる番号を選び〇で囲んでください。

1-まったくあてはまらない 2-あまりあてはまらない
3-どちらでもない 4-ややあてはまる 5-とてもよくあてはまる

自信と言語に関すること
1) 英語に対する自信不足 1 2 3 4 5
2) ライティング能力不足 1 2 3 4 5
3) リーディング能力不足 1 2 3 4 5
4) 英語で意見を表明する不安 1 2 3 4 5
5) 英語を使った時に他者からの反応に対してより敏感になること 1 2 3 4 5
6) 知らない専門用語 1 2 3 4 5
7) 英語使用に対するサポート不足 1 2 3 4 5
8) 書き言葉での投稿に対する自信不足 1 2 3 4 5
9) すぐに反応しなければならないことに対する自信不足 1 2 3 4 5
10) 現在の課題に対する自信不足 1 2 3 4 5
11) 難しすぎる課題 1 2 3 4 5
12) 振り返りのための時間不足 1 2 3 4 5

学習環境の設計に関すること
13) グループ作業でのはっきりとした見通しの無さ 1 2 3 4 5
14) 不明瞭な評価基準 1 2 3 4 5
15) 不適当な課題 1 2 3 4 5
16) 教員からのタイミングの良いサポート不足 1 2 3 4 5
17) 不明瞭な協働作業に対する指示 1 2 3 4 5
18) 教員からのオンライン会議における さまざまな場面での適切なサポート不足 1 2 3 4 5
19) 組織的な組み立てが不足した協働作業 1 2 3 4 5

テクノロジーに関すること
20) テクニカルサポート不足 1 2 3 4 5
21) オンラインツール（フォーラム、wiki など）の技術的な問題 1 2 3 4 5
22) 技術使用に関する多すぎるルール 1 2 3 4 5
23) オンラインツールを使いこなす時の難しさ 1 2 3 4 5
24) 技術使用に関するルール不足 1 2 3 4 5
25) オンライン学習中の技術的な問題に関する不安 1 2 3 4 5
（例: コンピュータの故障, ネット不調など）

協働作業に関すること
26) オンラインで他者との関係構築の難しさ 1 2 3 4 5
27) グループでの意志決定の難しさ 1 2 3 4 5
28) 協働作業に対するグループからのプレッシャー 1 2 3 4 5
29) オンライン会議で他の学習者の反応を待つのが嫌なこと 1 2 3 4 5
30) 協働作業で他者の投稿に対する信頼感の低さ 1 2 3 4 5
31) オンラインで学習するのが嫌なこと 1 2 3 4 5
32) 自分の学習スタイルとの不一致 1 2 3 4 5

その他にストレスを引き起こす要因はありますか？
・
・
コメント

Appendix B: Reading and writing self-efficacy scale

Translated from Shaw, E.J. (2007). The reading and writing self-efficacy beliefs of students with discrepant reading and writing performance. Unpublished Doctoral Dissertation, Fordham University.

以下のそれぞれの項目について英語を使って正確に読むことに対して、
0(全く自信が無い)から 100 (完璧に自信がある)までの範囲で、あなたはどれくらい自信があるかお答え下さい。0 から 100 の間のどの数で答えても構いません。

0 10 20 30 40 50 60 70 80 90 100

全く自信が無い　　完璧に自信がある

1. _____ 友人や家族からの手紙
2. _____ 食事のレシピ
3. _____ アパートの賃貸契約書
4. _____ 車両保険の契約書
5. _____ 就職用の履歴書
6. _____ コンピュータ操作のマニュアル
7. _____ 職務・社則に関する社員の服務書
8. _____ 選択肢問題で出題される高校の試験
9. _____ 高校レベルの得意科目の教科書
10. _____ 大学レベルの得意科目の教科書
11. _____ 得意科目に関連する技術書
12. _____ 新聞
13. _____ タイム や ニューズウィークなどの雑誌
14. _____ 短編小説
15. _____ 400 ページ程度の小説
16. _____ シェークスピアの戯曲
17. _____ 詩集
18. _____ 哲学書

0（全く自信なし）から 100（完璧に自信あり）までの段階で、以下の英語スキルについてどの程度自信があるか答えて下さい。0 から 100 までの好きな数字を選んで入れて記入して下さい。

1. _____ 英語の文字を認識する
2. _____ 個々の単語を発音する
3. _____ それぞれの品詞を認識する
4. _____ 文法的な文を認識する
5. _____ 複数形、接頭辞、接尾辞の意味を理解する
6. _____ 単文や長文を理解する
7. _____ 新しい語を発音する
8. _____ 短い物語や文章の重要なポイントを認識する
9. _____ 新しい教材を理解するために知っている知識を使う
10. _____ 文中のアイディアを比べ、差を見つけられる
11. _____ 文の作者の見方について推測ができる

以下のそれぞれの項目について正確な英語を使って書くことに対して、
0(全く自信が無い)から 100 (完璧に自信がある)までの範囲で、あなたはどれくらい自信があるかお答え下さい。0 から 100 の間のどの数で答えても構いません。

0　10　20　30　40　50　60　70　80　90　100

全く自信が無い　　完璧に自信がある

1. _____ 友人や家族に手紙を書く
2. _____ トランプゲームの遊び方を箇条書きにする
3. _____ コンピュータの操作マニュアルを書く
4. _____ あなた自身の職歴や技能を履歴書にまとめる
5. _____ テスト問題に対して１～２文で解答を書く
6. _____ テスト問題に対して１～２ページの解答を作文する
7. _____ 15～20 ページのレポートを書く
8. _____ あなたの論点を支持する例を提示しながら、自身の論を表明する
9. _____ 出版するために得意分野に関連した技術書を書く
10. _____ 新聞の編集者に対して意見文を書く
11. _____ ニューズウィークなどの一般的な雑誌に記事を書く
12. _____ 物語を書く
13. _____ 100 ページ程度の小説を書く
14. _____ 好きな題材で詩を書く
15. _____ 役に立つ講義ノートをとる
16. _____ 他の著者の批評を書く

0（全く自信なし）から 100（完璧に自信あり）までの段階で、以下の英語スキルについてどの程度自信があるか答えて下さい。0 から 100 までの好きな数字を選んで入れて記入して下さい。

1. _____ 1 ページの文章の全ての単語のスペルを正確に書く
2. _____ 1 ページの文章に正確に句読点を入れて書く
3. _____ 品詞（名詞、動詞、形容詞など）を正確に使う
4. _____ 簡単な文に句読点を入れて文法的に正確に書く
5. _____ 複数、動詞の時制、接頭辞、接尾辞を正確に使う
6. _____ 句読点を使い、文法的な重文や複文などを書く
7. _____ テーマに沿って文章をまとめ、段落を構成する
8. _____ 全体をよく編成し（意図を順番に、効果的な移行など）レポートを書く
9. _____ あなたの意図を支持するために明確な例を使う

Appendix C: Worked example without prompts

Come to Matsumoto for your college

Masayuki Kudo

Choosing a college town is an important decision. Where we live can decide our future. Sometimes we need to take a break from school, other times we should focus on studying. During both times, our environment is important. Although many cities offer attractions, Matsumoto is the perfect college town for two reasons: the balance between attractions and seclusion.

First of all, attractions are features of a city that make life more fun. Matsumoto has great attractions. There are many museums, galleries, and historical places. Matsumoto Castle, also known as "Crow Castle," is stunning all year-round. Mountains surround the city and outdoor adventure is close. Last but not least, Matsumoto Yamaga F.C. is a professional soccer team based here. *As you can see, there are many ways to break from studying. Matsumoto rivals bigger cities in Japan because of these attractions.*

Then, Matsumoto is not the only city in Japan with great attractions. Surely, Tokyo has the most museums and galleries in all of Japan. Himeji Castle, close to Kobe, is the largest castle in the country. Sapporo is the capital of Hokkaido, an international destination for outdoor adventure. Fukuoka, Yokohama and Nagoya all have professional sports teams. *No doubt, the big cities offer more attraction than Matsumoto. This small town must have something the big cities don't have.*

Second, Matsumoto's unique resource is its seclusion. Seclusion is important so students can focus on studying. Matsumoto is a small city in the mountains. Clean air and water are necessary for healthy living. This is better than the big and smoggy cities. In Matsumoto, students will enjoy peace and quiet. The people of Matsumoto are kind and warm. In the bigger cities, everyone is too

busy to be kind. They will not be found this in the bigger cities. *When it's time to focus on schoolwork, Matsumoto is the place to be because of its seclusion.*

Finally, although many cities have attractions, a balance of attractions and seclusion makes Matsumoto the best city for college. Matsumoto's attractions make life interesting. The big cities in Japan offer more attractions but Matsumoto also gives the student seclusion. Peace and quiet is essential for studying. This balance is just right for the college student. Come to Matsumoto and experience the balance of excitement and quiet.

紫：トランジション
アンダーライン：トピックセンテンス
赤、青、緑：「例」・「理由」
イタリック：コンクルーディングセンテンス

Appendix D: Worked example without prompts

以下の例を使って課題作成に役立てて下さい。
なお、この情報は他のグループには見せないでください。

Come to Matsumoto for your college

Masayuki Kudo (ZZZZ-09-XXX)

Choosing a college town is an important decision. Where we live can decide our future. Sometimes we need to take a break from school, other times we should focus on studying. During both times, our environment is important. Although many cities offer attractions, Matsumoto is the perfect college town for two reasons: the balance between attractions and seclusion.

First of all, attractions are features of a city that make life more fun. Matsumoto has great attractions. There are many museums, galleries, and historical places. Matsumoto Castle, also known as "Crow Castle," is stunning all year-round. Mountains surround the city and outdoor adventure is close. Last but not least, Matsumoto Yamaga F.C. is a professional soccer team based here. *As you can see, there are many ways to break from studying. Matsumoto rivals bigger cities in Japan because of these attractions.*

Then, Matsumoto is not the only city in Japan with great attractions. Surely, Tokyo has the most museums and galleries in all of Japan. Himeji Castle, close to Kobe, is the largest castle in the country. Sapporo is the capital of Hokkaido, an international destination for outdoor adventure. Fukuoka, Yokohama and Nagoya all have professional sports teams. *No doubt, the big cities offer more attraction than Matsumoto. This small town must have something the big cities don't have.*

Second, Matsumoto's unique resource is its seclusion. Seclusion is important so students can focus on studying. Matsumoto is a small city in the mountains. Clean air and water are necessary for healthy living. This is better than the big and smoggy cities. In Matsumoto, students will enjoy peace and quiet. The people

of Matsumoto are kind and warm. In the bigger cities, everyone is too busy to be kind. They will not be found this in the bigger cities. *When it's time to focus on schoolwork, Matsumoto is the place to be because of its seclusion.*

Finally, although many cities have attractions, a balance of attractions and seclusion makes Matsumoto the best city for college. Matsumoto's attractions make life interesting. The big cities in Japan offer more attractions but Matsumoto also gives the student seclusion. Peace and quiet is essential for studying. This balance is just right for the college student. Come to Matsumoto and experience the balance of excitement and quiet.

紫：トランジション
アンダーライン：トピックセンテンス
赤、青、緑：「例」・「理由」
イタリック：コンクルーディングセンテンス

☆課題を作成するにあたって、以下の質問を元に上の例文を参考にしてください。

1. この文章のタイトルは何ですか？
2. 各段落のトピックセンテンスは何ですか？ どんな情報が与えられていますか？
3. 各段落の文章にはサポーティングセンテンスはいくつありますか？
   それらは何を説明していますか？トランジションはありますか？
4. この文章で「理由」はいくつ提示されていますか？ 理由を示すのは、どれですか？
5. それぞれの「理由」には、いくつの「例」が提示されていますか？
   「例」が提示を示すのはどんな単語ですか？
6. この文章と各段落の結論文（コンクルーディングセンテンス）は何ですか？

☆最後に以下の３つの質問を頭に浮かべて自分の課題を見なおしてください。

a. 上の例に従って課題を作成するならば、自分たちのグループが作成した課題は、どのように情報を整理したら、もっと良い構成なるだろうか？
b. どのような理由や例が、高校生を説得するために説明、確認、対立的な手法を使って構成できるだろうか？
c. どのポイントで高校生を説得できているか？ またどれくらい説得しきれていないか？もう一度振り返り、確認しましょう。

Appendix E: the ESL composition profile

Adapted from Jacobs, H., Hartfiel, V. Hughey, J. and Wormuth, D. (1981). Testing ESL Composition: A. Jacobs, H (Ed). *Testing ESL composition: A practical approach.* Rowley, MA: Newbury House.

The extended profile criteria

THE ESL COMPOSITION PROFILE -- A GUIDE TO THE PRINCIPLES OF WRITING

6.2 The extended profile criteria

Since the criteria descriptors are only shorthand reminders of larger concepts in composition, a clear understanding of them is essential for effective use of the PROFILE. The concepts embody the essential principles of writing -- the rules, conventions,and guidelines -- that writers must observe to create a successful piece of writing. This section presents a detailed description of the concepts represented by the PROFILE criteria descriptors at the *Excellent to Very Good* mastery level. The other three levels of competence should be thought of as varying degrees of these extended criteria for excellent writing, with the primary distinguishing factor being the degree to which the writer's intended *meaning* is successfully delivered to the reader or is diminished or completely lost by insufficient mastery of the criteria for excellence. The PROFILE's first two mastery levels in each component (*Excellent to Very Good* and *Good to Average*) both indicate that successful communication has occurred (although differing in degree), whereas the two lower levels (*Fair to Poor* and *Very Poor*) suggest there is a communication breakdown of some sort -- either partial or complete. *Effect on meaning* thus becomes the chief criterion for distinguishing the degree to which the writer has mastered the criteria for excellent writing.

CONTENT

30-27	**EXCELLENT TO VERY GOOD: knowledgeable*substantive*thorough development of thesis* relevant to assigned topic**
26-22	**GOOD TO AVERAGE: some knowledge of subject* adequate range* limited development of thesis* mostly relevant to topic, but lacks detail**
21-17	**FAIR TO POOR: limited knowledge of subject* little substance* inadequate development of topic**
16-13	**VERY POOR: does not show knowledge of subject* non-substantive* not pertinent * OR not enough to evaluate**

DESCRIPTOR	CRITERIA
Knowledgeable	Is there understanding of the subject? Are facts or other pertinent information used? Is there recognition of several aspects of the subject? Are the interrelationships of these aspects shown?
Substantive	Are several main points discussed? Is there sufficient detail? Is there originality with concrete details to illustrate, define, compare, or contrast factual information supporting the thesis?
Thorough development of thesis	Is the thesis expanded enough to convey a sense of completeness? Is there a specific method of

	development (such as comparison/contrast, illustration, definition, example, description, fact, or personal experience)?
Relevant to assigned topic	Is all information clearly pertinent to the topic? Is extraneous material excluded?

ORGANIZATION

20-18	**EXCELLENT TO VERY GOOD: fluent expression* ideas clearly stated/supported* succinct*well-organized*logical sequencing*cohesive**
17-14	**GOOD TO AVERAGE: somewhat choppy*loosely organized but main ideas stand out*limited support* logical but incomplete sequencing**
13-10	**FAIR TO POOR: non-fluent* ideas confused or disconnected* lacks logical sequencing and development**
9-7	**VERY POOR: does not communicate* no organization*OR not enough to evaluate**

DESCRIPTOR	**CRITERIA**
Fluent expression	Do the ideas flow, building on one another? Are there introductory and concluding paragraphs? Are there effective transition elements -- words, phrases, or sentences -- which link and move ideas both within and between paragraphs?
Ideas clearly stated/supported	Is there a clearly stated controlling idea or central focus to the paper (a thesis)? do topic sentences in each paragraph support, limit, and direct the thesis?
Succinct	Are all ideas directed concisely to the central focus of the paper, without digression?
Well-organized	Is the overall relationship of ideas within and between paragraphs clearly indicated? Is there a beginning, a middle, and an end to the paper?
Logical sequencing	Are the points logically developed, using a particular sequence such as time order, space order, or importance? Is this development indicated by appropriate transitional markers?
Cohesive	Does each paragraph reflect a single purpose? Do the paragraphs form a unified paper?

VOCABULARY

20-18	**EXCELLENT TO VERY GOOD: sophisticated range* effective word/idiom choice and usage* word form mastery * appropriate register**
17-14	**GOOD TO AVERAGE: adequate range* occasional errors of word/idiom form, choice, usage *but meaning not obscured***
13-10	**FAIR TO POOR: limited range* frequent errors of word/idiom form, choice, usage* *meaning confused or obscured***
9-7	**VERY POOR: essentially translation* little knowledge of English vocabulary, idioms, word form* OR not enough to evaluate**

DESCRIPTOR	CRITERIA
Sophisticated range	Is there facility with words and idioms: to convey intended information, attitudes, feelings? to distinguish subtleties among ideas and intentions? to convey shades and differences of meaning? to express the logic of ideas? Is the arrangement and interrelationship of words sufficiently varied?
Effective word/idiom choice and usage	In the context in which it is used, is the choice of vocabulary accurate? idiomatic? effective? concise? Are strong active verbs and verbals used where possible? Are phrasal and prepositional idioms correct? Do they convey the intended meaning? Does word placement give the intended message? emphasis? Is there an understanding of synonyms? antonyms? homonyms? Are denotative and connotative meanings distinguished? Is there effective repetition of key words and phrases? do transition elements mark shifts in thought? pace? emphasis? tone?
Word form mastery	Are prefixes, suffixes, roots, and compounds used accurately and effectively? Are words correctly distinguished as to their function (noun, verb, adjective, adverb)?
Appropriate register	Is the vocabulary appropriate to the topic? to the audience? to the tone of the paper? to the method of development? Is the vocabulary familiar to the audience? Does the vocabulary make the intended impression?

LANGUAGE USE

25-22	**EXCELLENT TO VERY GOOD: effective complex constructions* few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions**
21-18	**GOOD TO AVERAGE: effective but simple constructions* minor problems in complex constructions * several errors of agreement, tense, number, word order/function, articles, pronouns, prepositions *but meaning seldom obscured***
17-11	**FAIR TO POOR: major problems in simple/complex constructions* frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions and/or fragments, run-ons, deletions * *meaning confused or obscured***
10-5	**VERY POOR: virtually no mastery of sentence construction rules* dominated by errors* does not communicate* OR not enough to evaluate**

DESCRIPTOR	**CRITERIA**
Effective complex constructions	Are sentences well-formed and complete, with appropriate complements? Are single-word modifiers appropriate to function? Are they properly formed, placed, sequenced? Are phrases and clauses appropriate to function? complete? properly placed? Are introductory *It* and *There* used correctly to begin sentences and clauses? Are main and subordinate ideas carefully distinguished? Are coordinate and subordinate elements linked to other elements with appropriate conjunctions, adverbials, relative pronouns, or punctuation? Are sentence types and length varied? Are elements parallel? Are techniques of substitution, repetition, and deletion use effectively?
Agreement	Is there basic agreement between sentence elements: auxiliary and verb? subject and verb? pronoun and antecedent? adjective and noun? nouns and quantifiers?
Tense	Are verb tenses correct? properly sequenced? Do modals convey intended meaning? time?

Number	Do nouns, pronouns, and verbs convey intended quality?
Word order/function	Is normal word order followed except for special emphasis? Is each word, phrase, and clause suited to its intended function?
Articles	Are *a, an,* and *the* used correctly?
Pronouns	Do pronouns reflect appropriate person? gender? number? function? referent?
Prepositions	Are prepositions chosen carefully to introduce modifying elements? Is the intended meaning conveyed?

MECHANICS:

5	**EXCELLENT TO VERY GOOD: demonstrates mastery of conventions* few errors of spelling, punctuation, capitalization, paragraphing**
4	**GOOD TO AVERAGE: occasional errors of spelling, punctuation, capitalization, paragraphing *but meaning not obscured***
3	**FAIR TO POOR: frequent errors of spelling, punctuation, capitalization, paragraphing * poor handwriting* *meaning confused or obscured***
2	**VERY POOR: no mastery of conventions* dominated by errors of spelling, punctuation, capitalization, paragraphing* handwriting illegible* OR not enough to evaluate**

DESCRIPTOR	**CRITERIA**
Spelling	Are word spelled correctly?
Punctuation	Are periods, commas, semicolons, dashes, and question marks used correctly? Are words divided correctly at the end of lines?
Capitalization	Are capital letters used where necessary and appropriate?
Paragraphing	Are paragraphs indented to indicate when one sequence of thought ends and another begins?
Handwriting	Is handwriting easy to read, without impeding communication?

Effects of Worked Examples on Stress, Cognitive Loads, and Performance in Online Collaboration

オンライン協働におけるワークトエグザンプルの ストレス、認知負荷、パフォーマンスへの影響

A Dissertation Presented to the Division of Education,
the Graduate School of International Christian University,
for the Degree of Doctor of Philosophy

国際基督教大学　大学院

November 21, 2013

Kudo, Masayuki
工藤　雅之