Appendix 1: Detailed protocol from TA-1 (CM-only and CM+CE groups)

Week 1:

CM-only (section 1 and 2): The students were given the pre-conceptual inventory test for the first 20 minutes. The next 5-10 minutes were then spent covering concept maps and how to approach them. An example of how to build a concept map was briefly covered which focused on phases of matter/gas laws. Students were then given a worksheet to work on for the remainder of the time. Answers for the review worksheet were covered at the end.

CM+CE (section 3 and 4): The students were given the pre-conceptual inventory test for the first 20 minutes. The next 5-10 minutes were spent going over concept maps and how to approach them. An example of how to build a concept map was briefly covered which focused on phases of matter/gas laws. 10-15 minutes were then spent showing the students how to complete the CE assignments. Five minutes were set aside to have students work on the first CE assignment, and 2 minutes were given for students to discuss with a partner the concepts they put down and any additional concepts they would like to add to their existing one or correct any concepts they had that may have an error. Instructions for the CE assignment were: 2 concepts for equilibrium/new material and 4 concepts for previous material. Students were instructed to draw a line across the page to separate their own ideas from those that they and their partner came up with. This CE was collected and scored that week by 2 TAs and 3 undergraduate student researchers. Any remaining time was used to start the review worksheet.

Week 2:

CM-only (section 1 and 2): Students were given a worksheet to work on for ~35 minutes. Answers for the worksheet were covered at the end.

CM+CE (section 3 and 4): The first 5 minutes of the class focused on hitting some of the major issues that were seen from the first week of scoring the CEs. Students were then instructed to work on the new CE given to them but this time they need 10 total topics (4 from current equilibrium chapter and the rest from anything else prior but relevant to the topic). Students were given 10 minutes to work on this individually and then instructed to work with one other person for 5 minutes to discuss their concepts. Again, students would draw a line across their page to separate their independent concepts from that of they and their partner as well as correct any concepts they had that when discussed they deemed to be incorrect/not relevant/not distinct. This CE was collected and scored that week by 2 TAs and 3 undergraduate student researchers. The remainder of the time was set aside to complete and discuss the weekly activity. The week 1 CE was handed back.

Week 3:

CM-only (section 1 and 2): Students were given a worksheet to work on for ~35 minutes. Answers for the worksheet were covered at the end.

CM+CE (section 3 and 4): The first 5 minutes of the class focused on showing the students a completed CE that was put together by myself and another undergraduate researcher. We showed

them examples of the different concepts they could have talked about, especially those that were more obscure, like electron configuration, to get them thinking about what they could potentially talk about. Students were then instructed to work on the new CE given to them and identify 10 total concepts they could discuss (4 from current acid/base chapter and the rest from anything else prior but relevant to the topic). Students were given 15 minutes to work on this individually and then instructed to work with one other person for 5 minutes to discuss their concepts. Again, students would draw a line across their page to separate their independent concepts from that of they and their partner as well as correct any concepts they had that when discussed they deemed to be incorrect/not relevant/not distinct. This CE was collected and scored that week by 2 TAs and 3 undergraduate student researchers. The remainder of the time was set aside to complete and discuss the weekly activity. The week 2 CE was handed back.

Week 4:

CM-only (section 1 and 2): Students were given a worksheet to work on for ~35 minutes. Answers for the worksheet were covered at the end.

CM+CE (section 3 and 4): Students were instructed to work on the new CE given to them and identify 10 total concepts they could discuss (4 from current acid/base chapter and the rest from anything else prior but relevant to the topic). Students were given 15 minutes to work on this individually and then instructed to work with one other person for 5 minutes to discuss their concepts. Again, students would draw a line across their page to separate their independent concepts from that of they and their partner as well as correct any concepts they had that when discussed they deemed to be incorrect/not relevant/not distinct. This CE was collected and scored that week by 2 TAs and 3 undergraduate student researchers. The remainder of the time was set aside to complete and discuss the weekly activity. The week 3 CE was handed back.

Appendix 2: Detailed protocol from TA-2 (CE-only and control group)

Week 1:

Control group (section 5 and 6): I introduced myself to the students, covering my name, contact info, and where I am in my graduate education. We took a short poll where I asked students what they were studying and if the class was part of a learning community. After introductions and questions, the students were asked to log into the learning management system iLearn and take the pre-conceptual inventory test for approximately 20 minutes. Most students took less than the allotted time, and I closed the online link after all students had submitted their tests. After the test, I reassured the students that it was only their participation that was counting for their grade, as several were concerned about seeing/not seeing their scores. Students were then given a worksheet covering concepts from the last quarter of general chemistry, and we spent a few minutes at the end going over answers. Students were let out of discussion as much as ten minutes early.

CE-only (section 7 and 8): I introduced myself to the students, covering my name, contact info, and where I am in my graduate education. We took a short poll where I asked students what they were studying and if the class was part of a learning community. After introductions and

questions, the students were asked to log into iLearn and take the pre-conceptual inventory test for approximately 20 minutes. Most students took less than the allotted time, and I closed the online link after all students had submitted their tests. After the test, I reassured the students that it was only their participation that was counting for their grade, as several were concerned about seeing/not seeing their scores. The next 10-15 minutes were spent using a PowerPoint to introduce the creative exercises, and explain what counted as correct, distinct, and relevant. The students were then given the first CE assignment and the students were given 10 minutes to work on them on their own, then a few minutes to trade with a partner and "grade" or give feedback, with the last minutes of class for them to talk to their partners and surrounding classmates. During this time the students were encouraged to add to their list any statements from their classmates that they liked, or any new ideas to the bottom of their paper. The CE papers were collected and scored by the 2 TAs and 3 undergraduate student researchers. The students were given access to the first week's worksheet, and it was also available on iLearn.

Week 2:

Control group (section 5 and 6): We spent the first ten minutes talking about class logistics and any questions the students had about current material. No major class content was covered. Students were then given a hard copy of the week's worksheet and worked on it for most of the class time. We covered several of the answers at the end of class, and the students could leave early.

CE-only (section 7 and 8): The first ten minutes of class were spent covering student questions and logistics, with a PowerPoint covering the major issues seen in the grading of the first weeks CE assignment. After highlighting these issues, students were given the same CE paper and given approximately 10 minutes to come up with a full 10 statements, four of which were supposed to be from current equilibrium material. They then traded papers with a partner and worked together for around 10-15 minutes, where grading of their peers' papers as well as adding new statements was encouraged. The CE was then scored by the same people as week 1, and any remaining time was given to the worksheet, though no completion was required. The week 1 CE was handed back at the end of the week 2 CE assignment.

Week 3:

Control group (section 5 and 6): We spent the first ten minutes talking about class logistics and any questions the students had about current material. Students were then given a hard copy of the weeks worksheet and worked on it for most of the class time. When students had a question about a specific problem on the worksheet, we talked about it as a class. We covered several of the answers at the end of class.

CE-only (section 7 and 8): The first 5 minutes of the class focused on showing the students a completed CE that was put together by myself and an undergraduate student. We showed them examples of the different concepts they could have talked about, especially those that were more obscure, like electron configuration, to get them thinking about what they could potentially talk about. Students were then instructed to work on the new CE given to them and identify 10 total concepts they could discuss (4 from current acid/base chapter and the rest from anything else prior but relevant to the topic). Students were given 15 minutes to work on this individually and

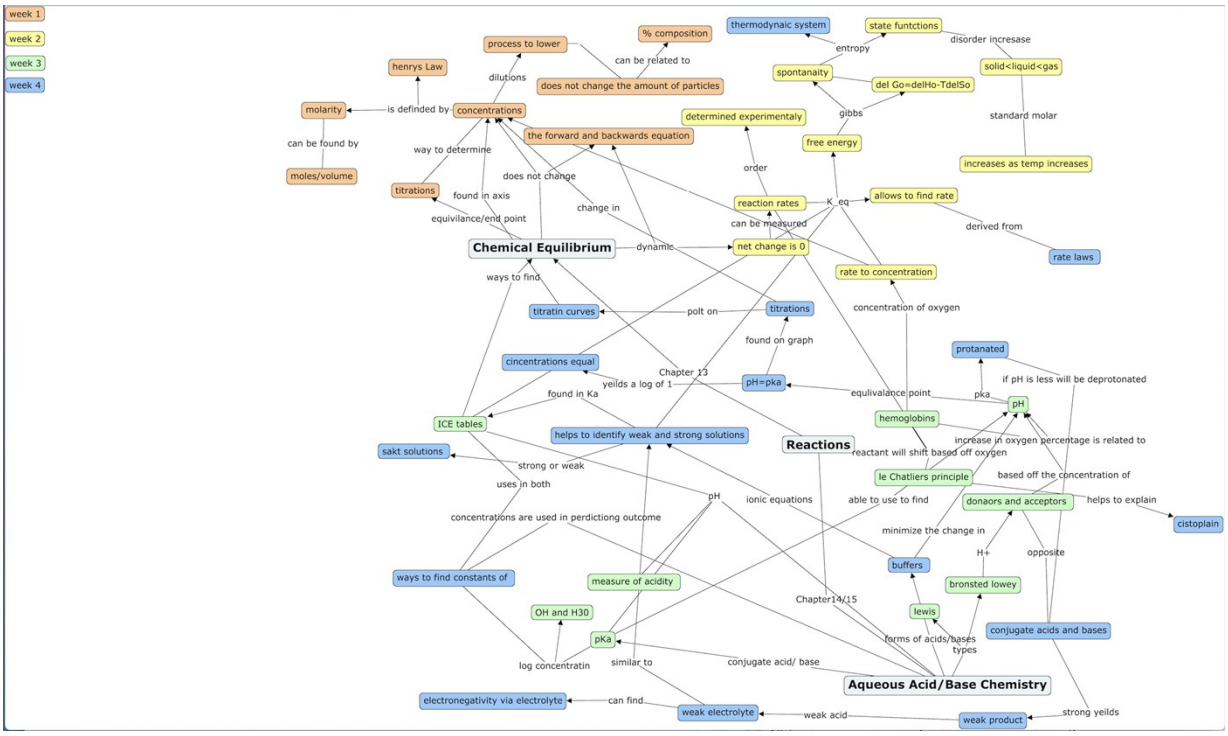
then instructed to work with one other person for 5 minutes to discuss their concepts. Again, students would draw a line across their page to separate their independent concepts from that of they and their partner as well as correct any concepts they had that when discussed they deemed to be incorrect/not relevant/not distinct. This CE was collected and scored that week by 2 TAs and 3 undergraduate student researchers. The remainder of the time was set aside to complete and discuss the weekly activity. The week 2 CE was handed back.

Week 4:

Control group (section 5 and 6): Students were given a worksheet to work on for ~35 minutes. Answers for the worksheet were covered at the end

CE-only (section 7 and 8): Students were instructed to work on the new CE given to them and identify 10 total concepts they could discuss (4 from current acid/base chapter and the rest from anything else prior but relevant to the topic). Students were given 15 minutes to work on this individually and then instructed to work with one other person for 5 minutes to discuss their concepts. Again, students would draw a line across their page to separate their independent concepts from that of they and their partner as well as correct any concepts they had that when discussed they deemed to be incorrect/not relevant/not distinct. This CE was collected and scored that week by 2 TAs and 3 undergraduate student researchers. The remainder of the time was set aside to complete and discuss the weekly activity. The week 3 CE was handed back.

Appendix 3: Example concept map created by a student in the study



Appendix 4: Creative Exercises used in the study

Week 1 & week 2 (recitation sections): CE prompt for the chemical equilibrium unit

Based on everything you learned in General Chemistry, write down as many correct, distinct and relevant statements you can about the following prompt in bold.

Ammonia is a crucial component of many fertilizers. The Haber process (see reaction below) is used to make the fertilizer needed in agriculture and sustains 40% of the world's population. In the manufacture of ammonia, conditions of high temperature, high pressure, and the presence of an iron catalyst are used.

$$\mathbf{N_2\,(g) + H_2\,(g) \rightleftharpoons NH_3\,(g)}$$

At a certain temperature, the equilibrium constant for this reaction is 6.00×10^{-2} M^{-2}. The equilibrium concentrations of nitrogen and ammonia are 0.080 M and 0.13 M, respectively, and the reaction is taking place in a 5.00 L container.

You'll receive 2 points for each statement. Ten statements will get you full credit for the problem. At least **four** statements need to be from the content you learned in Chapter 13 (chemical equilibrium).

Example student responses:

1. *The balanced reaction is: N_2 (g) + 3 H_2 (g) ⇌ 2 NH_3 (g)*
2. *The Kc <1 which means there are more reactants than products*
3. *Increasing the pressure would shift the equilibrium to the right [products]*
4. *The iron catalyst used does not affect the [equilibrium] concentrations of both reactants and products*
5. *The equilibrium concentration of hydrogen can be calculated:*
   $6.00 \times 10^{-2} = [0.13]^2 / [0.08] [x]^3$
   x = 1.52 M

Week 3 & week 4 (recitation sections): CE prompt for the acid-base unit

Based on everything you learned in General Chemistry, write down as many correct, distinct and relevant statements you can about the following prompt in bold:

Heartburn occurs when the stomach produces too much acid. Hydrochloric acid is the main substance that causes heartburn. Antacids contain active ingredients (for example, sodium bicarbonate) to neutralize the hydrochloric acid in the stomach and relieve heartburn. Consider 125 mL of 2.50 M hydrochloric acid reacting with 5.00 grams of sodium bicarbonate. This reaction produces 37.34 kJ of heat.

You'll receive 2 points for each statement. Ten statements will get you full credit for the problem. At least **four** statements need to be from the content you learned in Chapter 14 & 15 (acids/bases).

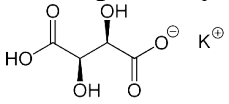
Example student responses:

1. *According to Brønsted-Lowry, HCl is an acid because it is a proton donor.*
2. *The pH of the HCl is -0.398*
3. *NaCl is the conjugate base of HCl*
4. *The electron configuration of chlorine is [Ne] 3s2 3p5*
5. *Increasing the temperature of the system will shift the reaction to the reactant side and decreasing the temptation will shift it to the product side.*

Week 5 (first midterm exam): CE prompt for both chemical equilibrium and acid-base unit

Based on everything you learned in General Chemistry, write down as many correct, distinct and relevant statements you can about the following prompt in bold.

Potassium bitartrate ($KHC_4H_4O_6$), or cream of tartar, is a byproduct of winemaking. It is a common component of baking powder and is also used to stabilize egg whites for whipping in the baking industry. When a saturated solution of potassium bitartrate is made at 25 °C, a pH of 3.557 is observed, and this solution can be used as a buffer solution. The pK_{a1} and pK_{a2} of tartaric acid are 3.22 and 4.85, respectively.

Structure of potassium bitartrate: 

You'll receive 2 points for each statement. Ten statements will get you full credit for the problem. At least **four** statements need to be from the content you learned in Chapter 13 (chemical equilibrium) and/or Chapter 14 & 15 (acids/bases).

Example student responses:

1. *The Kas for this reaction are both less than 1 so that means the reactants are favored.*
2. *Weak acids and bases result in reversible reactions because they do not dissociate completely.*
3. *Hydrogen bonding occurs in the structure of potassium bitartrate*
4. *The first ionization of tartaric acid yields a higher concentration of H_3O^+ than the second ionization.*
5. *$[H^+] = 10^{-3.557} = 0.000277\ M$*

Appendix 5: Item difficulty, and internal reliability of the combined concept inventory items used in the study.

Items 1-5 are from the *Journal of Chemical Education* Library of Conceptual Questions **(https://www.chemedx.org/JCEDLib/QBank/collection/CQandChP/CQs/LibraryCQ/EquilibriumCQ.html).**

Items 6-12 are from the acid-base inventory developed by Jensen **(**Jensen J.D., (2013), Students' understandings of acid-base reactions investigated through their classification schemes and the acid-base reactions concept inventory (Doctoral dissertation, Miami University). Retrieved from https://etd.ohiolink.edu/.).

Table 1: Item difficulty statistics for the post-test concept inventory questions.

Item Statistics			
	Mean	Std. Deviation	N
Item 1	0.4625	0.49963	240
Item 2	0.7875	0.40993	240
Item 3	0.9083	0.28916	240
Item 4	0.8542	0.35368	240
Item 5	0.9250	0.26394	240
Item 6	0.6333	0.48290	240
Item 7	0.5917	0.49255	240
Item 8	0.9667	0.17988	240
Item 9	0.9042	0.29498	240
Item 10	0.8708	0.33609	240
Item 11	0.5458	0.49894	240
Item 12	0.5125	0.50089	240

Table 2: Stratified alpha reliability coefficient for concept inventory.

Stratified Alpha (α_s) = 1- [Σ (variance of each dimension) (1- α_i)/variance of all items] α_s =0.566
Coefficient alpha for all test items = 0.554 (n = 12; variance of all test items = 3.87)
Test Item Concept Dimensions (α and variance of each dimension)
Items 2-5 (Equilibrium); α_i = 0.221; variance = 0.535
Items 1, 6-12, 10 (Equilibrium); α_i = 0.552; variance = 2.82

As described by: Widhiarso, W., & Ravand, H. (2014). Estimating reliability coefficient for multidimensional measures: A pedagogical illustration. *Review of Psychology*, *21*(2), 111–121.

Appendix 6 Focus group interview protocol and survey questions

Table 1: Focus group interview protocol

Part 1 Conceptual Questions:

1. Explain what it means for a chemical reaction to achieve equilibrium.
2. If a chemical reaction has achieved equilibrium, describe two or three attributes that would be associated with such a reaction.
3. If you increase the concentration of a reactant in a reaction that is at equilibrium, describe what will happen to the reaction (how will the reaction change?).
4. Describe the difference between a strong acid and a weak acid.
5. Draw the chemical reaction between HCl and NH_3 and explain how the molecules interact in this acid/base reaction; Use the Brønsted-Lowry definition of acids/bases to describe this reaction.
6. Draw the chemical reaction between HCl and NH_3 and explain how the molecules interact in this acid/base reaction; Use the Lewis definition of acids/bases to describe the above reaction.
7. Explain how the acetate buffer (mixture of CH_3COOH/CH_3COONa) is able to minimize the change in pH if dilute HCl is added to the buffer solution.
8. Explain how the acetate buffer (mixture of CH_3COOH/CH_3COONa) is able to minimize the change in pH if dilute NaOH is added to the buffer solution.

Part 2 Perception Questions:

1. Have you used CE/CM in your previous chemistry courses?
2. How do you like CE/CM in general as a chemistry assessment? Why?
3. Do you think CE/CM help you understand chemistry conceptually? Please elaborate the reason for your answer.
4. Do you think CE/CM help you make connections among chemistry concepts in General Chemistry? Please elaborate the reason for your answer.
5. How easy did you find to answer CE/ make CM?
6. What learning approaches did you use to prepare for CE/CM?Are there any differences when you prepared for this type of assessment versus other types of assessment such as multiple-choice questions? How do CE/CM impact your learning approaches in chemistry?
7. Do you believe you performed better in CE/CM than other types of chemistry assessment such as multiple-choice questions? Do you believe CE/CM help you improve your performance overall in this course?
8. Do you believe the use of CE/CM change the way you think about chemistry?
9. What suggestions do you have for your instructor or TA to improve CE/CM as a classroom assessment?
10. What suggestions do you have for future students in this course about how to improve their performance on CE/CM?
11. Would you like to see CE/CM in your future exams in this course or future chemistry courses? Why?
12. Any additional comments about CE/CM?
13. What are the similarities and differences between CE and CM for you (for CE+CM group only)? Do you believe the use of both assessments help you more for learning chemistry? Why?

Table 2: Survey questions

CM-only	CE-only	CE+CM
1. Do you think the Concept Maps help you understand chemistry conceptually: Yes or No? 2. Please explain why you chose the answer to question 1? 3. Do you think Concept Maps help you make connections among chemistry concepts: Yes or No? 4. Please explain why you chose the answer to question	1. Do you think Creative Exercises help you understand chemistry conceptually: Yes or No? 2. Please explain why you chose the answer to question 1? 3. Do you think Creative Exercises help you make connections among chemistry concepts: Yes or No? 4. Please explain why you chose the answer to question 3? 5. Please explain why you chose the answer to question 3?	1. Do you think Creative Exercises help you understand chemistry conceptually: Yes or No? 2. Please explain why you chose the answer to question 1? 3. Do you think Creative Exercises help you make connections among chemistry concepts: Yes or No? 4. Please explain why you chose the answer to question 3? 5. Do you think the Concept Maps help you understand chemistry conceptually: Yes or No 6. Please explain why you chose the answer to question 5? 7. Do you think Concept Maps help you make connections among chemistry concepts: Yes or No? 8. Please explain why you chose the answer to question 7? 9. Do you think doing the Creative Exercises helped prepare you for making better concept maps: Yes or No? 10. Please explain why you chose the answer to the question above?

Appendix 7: ANCOVA results and tests for assumptions; analysis in which the mean concept inventory post-test scores were compared between all four study groups (concept inventory pre-test held constant as a covariate).

Table 1: Full ANCOVA results

Tests of Between-Subjects Effects

Dependent Variable: concept inventory post-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	102.416[a]	4	25.604	7.207	0.000	0.112
Intercept	1234.633	1	1234.633	347.499	0.000	0.604
Pretest	92.434	1	92.434	26.016	0.000	0.102
GroupAll	9.711	3	3.237	0.911	0.436	0.012
Error	810.064	228	3.553			
Total	19588.000	233				
Corrected Total	912.481	232				

a. R Squared = .112 (Adjusted R Squared = .097)

Table 2: Adjusted means for the dependent variable (these are the adjusted means of the post-test dependent variable after the concept inventory pre-test covariate has been statistically controlled).

Estimates

Dependent Variable: concept inventory post-test

GroupAll	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Control	8.935[a]	0.246	8.451	9.420
Concept	8.752[a]	0.254	8.250	9.253
CE	8.828[a]	0.243	8.348	9.307
ConceptCE	9.285[a]	0.245	8.802	9.769

a. Covariates appearing in the model are evaluated at the following values: Pretest = 6.0858.

Table 3: Bonferroni-corrected between-groups pairwise comparisons (dependent variable = concept inventory post-test score). This provides the comparison of post-test dependent variable between each of the study groups, while correcting for family-wise error.

Pairwise Comparisons

Dependent Variable: concept inventory post-test

(I) GroupAll		Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
Control	Concept	0.184	0.354	1.000	-0.759	1.126
	CE	0.108	0.346	1.000	-0.813	1.029
	ConceptCE	-0.350	0.348	1.000	-1.276	0.575
Concept	Control	-0.184	0.354	1.000	-1.126	0.759
	CE	-0.076	0.352	1.000	-1.013	0.861
	ConceptCE	-0.534	0.353	0.793	-1.474	0.407
CE	Control	-0.108	0.346	1.000	-1.029	0.813
	Concept	0.076	0.352	1.000	-0.861	1.013
	ConceptCE	-0.458	0.346	1.000	-1.378	0.462
ConceptCE	Control	0.350	0.348	1.000	-0.575	1.276
	Concept	0.534	0.353	0.793	-0.407	1.474
	CE	0.458	0.346	1.000	-0.462	1.378

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Table 4: Levene's test of equality of error of variances (if $p > 0.05$, the null hypothesis that states the error of variance of the dependent variable is equal across the study groups cannot be rejected). The equality of error variances for the dependent variable across the study groups is a required assumption for ANCOVA analyses.

Levene's Test of Equality of Error Variances[a]

Dependent Variable: concept inventory post-test

F	df1	df2	Sig.
0.863	3	229	0.461

Table 5: Skewness and Kurtosis Analyses. In order to meet the assumptions for ANCOVA analyses, skewness should be between -1 and +1, kurtosis should be between -2 and +2, and the absolute value of skewness and kurtosis should be less than 3x the standard error.

Statistics

Concept inventory post-test		
N	Valid	233
	Missing	0
Skewness		-0.422
Std. Error of Skewness		0.159
Kurtosis		0.551
Std. Error of Kurtosis		0.318

Figure 1: Histogram of post-test score distribution with best-fit normal plot.

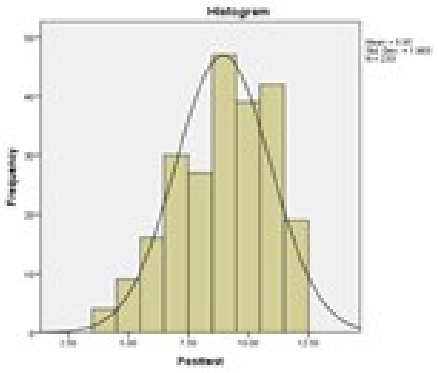


Table 6: Correlation between concept inventory pre- and post-test. The fact the correlation between pre- and post-test scores is significant indicates it is appropriate to include the concept inventory pre-test score as a covariate in the ANCOVA model.

Correlations

		Pretest	Posttest
Pretest	Pearson Correlation	1	.319**
	Sig. (2-tailed)		0.000
	N	233	233
Posttest	Pearson Correlation	.319**	1
	Sig. (2-tailed)	0.000	
	N	233	233

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix 8: ANCOVA results and tests for assumptions; analysis in which the mean CE scores were compared between the CE and CE+CM study groups (concept inventory pre-test held constant as a covariate).

Table 1: Full ANCOVA results.

Tests of Between-Subjects Effects

Dependent Variable: last CE scores

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	275.551[a]	2	137.775	10.672	0.000	0.155
Intercept	876.575	1	876.575	67.899	0.000	0.369
Pretest	262.851	1	262.851	20.360	0.000	0.149
GroupAll	14.670	1	14.670	1.136	0.289	0.010
Error	1497.559	116	12.910			
Total	21496.000	119				
Corrected Total	1773.109	118				

a. R Squared = .155 (Adjusted R Squared = .141)

Table 2: Adjusted means for the dependent variable (these are the adjusted means of the post-test dependent variable after the concept inventory pre-test covariate has been statistically controlled).

Estimates

Dependent Variable:

GroupAll	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
CE	12.526[a]	0.464	11.607	13.445
ConceptCE	13.228[a]	0.468	12.302	14.155

a. Covariates appearing in the model are evaluated at the following values: Pretest = 6.0168.

Table 3: Levene's test of equality of error of variances (if $p > 0.05$, the null hypothesis that states the error of variance of the dependent variable is equal across the study groups cannot be rejected). The equality of error variances for the dependent variable across the study groups is a required assumption for ANCOVA analyses.

Levene's Test of Equality of Error Variances[a]

Dependent Variable: last CE scores

F	df1	df2	Sig.
1.121	1	117	0.292

Table 4: Skewness and Kurtosis Analyses. In order to meet the assumptions for ANCOVA analyses, skewness should be between -1 and +1, kurtosis should be between -2 and +2, and the absolute value of skewness and kurtosis should be less than 3x the standard error.

Statistics

CEscore		
N	Valid	119
	Missing	0
Skewness		-0.393
Std. Error of Skewness		0.222
Kurtosis		-0.229
Std. Error of Kurtosis		0.440

Figure 1: Histogram of post-test score distribution with best-fit normal plot.

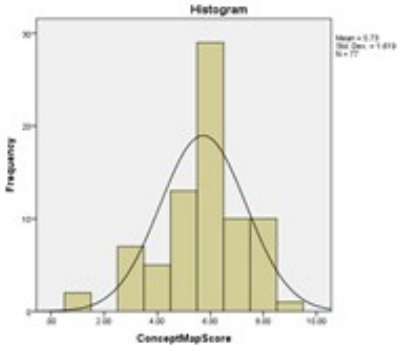


Table 5: Correlation between concept inventory pre-test and final CE scores. The fact the correlation between pre-test scores and final CE scores is significant indicates it is appropriate to include the concept inventory pre-test score as a covariate in the ANCOVA model.

Correlations

		Pretest	CEscore
Pretest	Pearson Correlation	1	.384**
	Sig. (2-tailed)		0.000
	N	233	119
CEscore	Pearson Correlation	.384**	1
	Sig. (2-tailed)	0.000	
	N	119	119

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix 9: Representative quotations of the codes unique to creative exercises or concept maps alone

Table 1:Representative quotations of the codes unique to concept maps alone

Code	Representative Quotation
Word bank	…it's better to have a word bank to show all the possible terms you can use.
Personalization	I pretty much color code the main ideas, so like the big boxes like. I'd like, I color code them as like red, or as like filling off of the bigger idea, so like the smaller boxes, I would like color like code green or other colors. But I pretty much just color code to be able to differentiate between the main idea and what's like a concept of main, the main idea.
Individuality	I do believe that looking at other people's concept maps probably would not help me that much because I think the whole point of a concept map is not for like a grade, but just doing it for yourself.
Learning curve	I would tell future students to like not be afraid to ask questions if they are not sure on how to like uh- how something works on a concept maps because for me I kind of had trouble like getting used to the software at first, and that kind of like lagged me behind, so I was able to like make my c-map right away so- I would highly recommend students to like ask for help if they do not, completely know how to use the software.
Mathematical equations and mathematical symbols	I believe that concept maps do help me understand chemistry conceptually because pretty much, concept maps, you can't really lay out- I mean you can lay out equations on there, but you can't really like doing the math so you are forced to like lay out the concepts and actually think about what you are talking about.
Studying consistently	I definitely- even if my professor does not require concept maps I definitely think I would use it in the future, because it is such a helpful study tool, and it has helped me conceptualize a lot of concepts in terms of chemistry.
Future exams	I think it is very helpful because you pretty much lay out all the concepts and then you are pretty much building off of the concepts and adding new things and so if you need to go back like, we just took our first midterm, so like we are about to prepare for our second midterm, and then for our final we can just look back at the concept map and look at all the concepts we listed out and then just study off

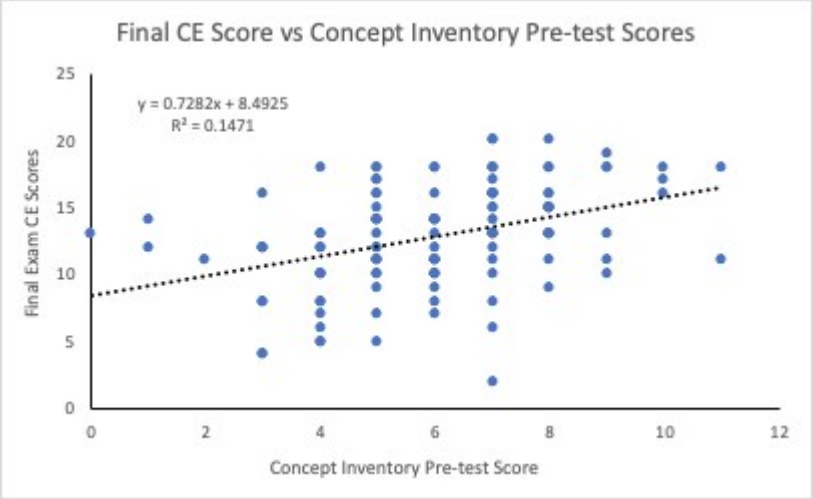
	of those concepts and just build off of that.
Organizing concepts	And I would just like- for me I would just take my notes on the concept map, like just linking all those ideas together and then connecting them to other chapters and things that I have learned from previous classes I just put it all in that map so I just have everything in one place. I thought it was a really nice way to organize all of that.

Table 2: Representative quotations of the codes unique to creative exercises alone

Code	Representative Quotation
Specific context	I prefer the creative exercises, just because it gives me an actual example that I can think of and then apply my concepts to.
Group collaboration	Because it can help you connect ideas together, maybe if you did it in groups then you could bounce ideas off each other.
hints	You have some kind of guide to help you come up with those statements then maybe that would be more useful.
Criteria	And also, as I come up with my statements, I'm not sure if they're concise, relevant, and what was it? Like if they meet the three qualifications.
Expectations	I think the guidelines are a little like, in a grey area.
Constructive feedback	And when I looked at what points I got missed, there were just a bunch of half, half, half, half, half, and I don't know why I got half off, because you know- Yeah the grade system they used was (inaudible) so they can, you know, they can see how they graded. But there's no feedback, so that's, that's the main thing- I don't know what I did wrong, I don't know why it's half or why it's not- That's why it's hard to quantitatively show what, like what is an answer for the creative exercises but for the free response question and the multiple choice, I can see what I wrote, what went wrong and I can see the process…
Current topics	I just thought of like four that are from last lecture, then six random ones that I happen to remember.

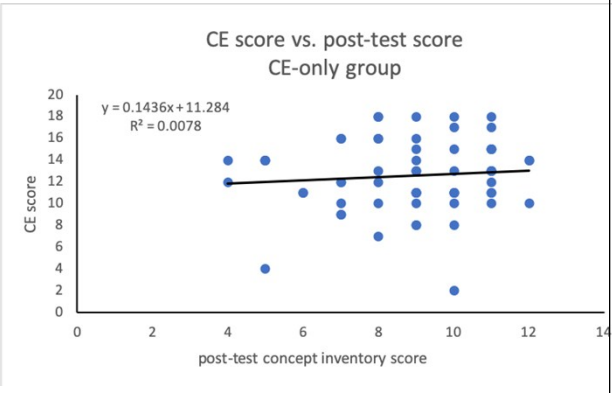
Appendix 10: Scatter plots between concept inventory pre and post-test scores and CE scores

Correlation and scatter plot for concept inventory pre-test scores and CE scores

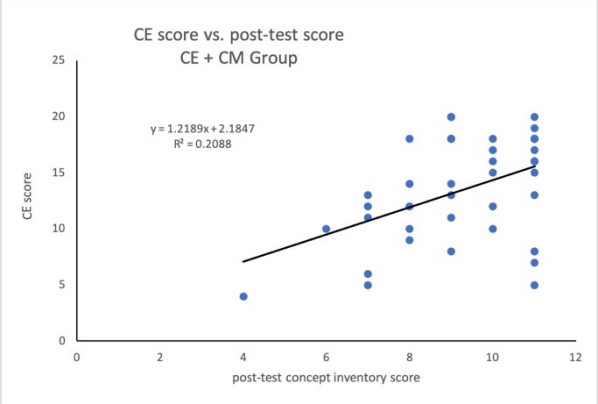


Pearson's r = 0.384 ; p < 0.001

Correlations and scatter plots between concept inventory post-test scores and CE scores (comparison between CE-only group and CM+CE group)



Pearson's r = 0.089; p = 0.501



Pearson's r = 0.457; p = 0.003