Chapter 27
Gamification in General Chemistry

Dave Allen Jenkins and Diana Mason

Games are very compelling (ESA 2012) and students often spend hours engaged in a single game (Rideout 2011; Rideout et al. 2010; Gutnick et al. 2011). Humans learn by playing, and one of the best ways to learn is through failure in low-stress situations, which can be introduced in an academic environment (Toprac 2018). Games have been widely explored as a tool to enhance learning (Civic Enterprises 2006); however research on gamification has focused primarily on preschool and elementary-age children in particular, and university-level studies related to academic success in STEM areas are few. We agree with the suggestion of Nobel Laureate Carl Wieman (2014) that educational research could be employed to compare different active learning methods. In this chapter we describe a study to assess the effects of a computer-based gaming environment developed by the first author, called *LABMATTER*. The game was used as a supplement to enhance students' knowledge base in first-semester general chemistry and was compared to another active learning control group and a group of students who experienced only traditional pedagogy without a High-Impact Practice (HIP) intervention.

Theoretical Prospectus

Research into game-based approaches for adults is far less common than games developed for younger players. Educational games and the concept of play are very difficult to define as they depend on several transient, contradictory, and subject-dependent qualities. Play is often described as the opposite of serious work. The

D. A. Jenkins
Physical Science, Tarrant County College, Hurst, TX, USA

D. Mason (✉)
Department of Chemistry, University of North Texas, Denton, TX, USA
e-mail: Diana.Mason@unt.edu

J. J. Mintzes, E. M. Walter (eds.), *Active Learning in College Science*,
https://doi.org/10.1007/978-3-030-33600-4_27

boundaries are often bent, broken, or blurred during play, and, as such, it can be challenging to differentiate between work and play.

The definitions of what establishes game and play are broad. Play is perceived as being beneficial in the promotion and development of imagination, creativity, and spontaneous learning. These same characteristics are often necessary for a student's growth in any educational setting. The utility of approaches that involve gaming or play in education may result from tactile activity with physical and emotional sensations. Creative skills are best refined using a range of approaches so that students can investigate issues from different perspectives. Several researchers suggest that the fundamental motivation for all games is to learn, and play is a safe way to learn, because games are effective supplements to traditional classroom pedagogies at all levels (Clark et al. 2010; Clark and Sheridan 2010). Educational methods based on games are often founded in a constructivist theory of learning. One requirement of constructivism is that students are challenged; specifically they are called to engage prior knowledge with the knowledge intended to be learned. New knowledge is constructed from interactions of people with materials, not simply transmitted (Williamson 2008). Challenges enable each student to devise his or her own knowledge and may also require alternative teaching approaches to engage and motivate the novice.

Recent efforts have been made in using games to encourage construction of scientific knowledge (Clark et al. 2010). Four fundamental constructs within the *unified learning model (ULM)* address the conditions required for meaningful learning; these include prior knowledge, engagement, working memory allocation, and motivation (Shell et al. 2010) (Fig. 27.1). Working memory, where connections are made, is governed by motivation and meaningful engagement and is the site of assimilation of new material. Accordingly, the storage of learned information is affected by prior knowledge and determines one's ease, speed, and efficiency of information processing.

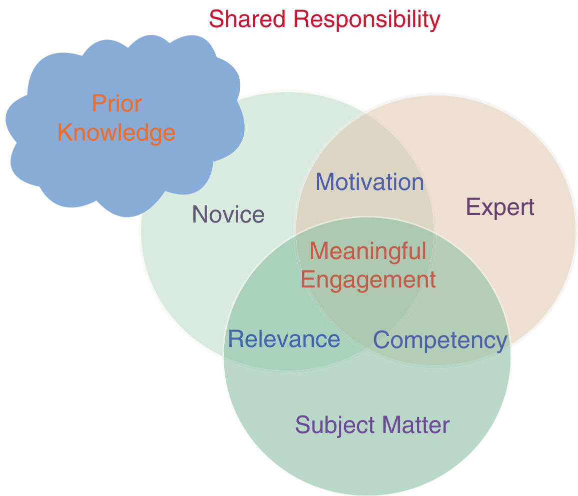

Fig. 27.1 Aspects of the unified learning model (ULM)

In the context of digital games, specifically first-person role-playing games, scaffolding is a particularly useful component for supporting individuals of varying ability levels. In educational games, scaffolding typically takes the form of *leveling* and *hints or clues*. Computer games most often include a sequence of game levels that increase in difficulty as the player becomes increasingly more efficient at each level and advances in the game. Educational learning increases by elevating the cognitive demands during game progression, rather than simply the use of psychomotor skills (eye-hand coordination and manual dexterity). Accordingly, beginning levels provide more scaffolding for students' learning, which is gradually alleviated as the player completes each game level. The player's selection of a wrong answer also provides an opportunity for additional support or tutorials. When students answer incorrectly, the game can provide additional support in the form of hints or clues for the approaching problem. These hints can offer more help with each subsequent problem, until the student can solve problems beyond the novice level. Clues prompt an accelerated retrieval of sensory data to which humans already receive a tremendous amount of sensory input at any given moment. Sensory allotments simply collect pieces of data into an output that is transferred to the appropriate parts of the brain.

Development

Gamification using an escape room methodology and activities is in its infancy in formal educational settings at the university level. (In an "escape room" game, people are trapped inside a space for a specific amount of time and need to solve a number of puzzles or problems to get out.) A computer game's "interface" is the system of hardware (e.g., joystick, mouse, touchscreen) and software applications (e.g., move, scroll, or drag) that are used to access and interact with the game. The functions associated with the interface can serve to either facilitate or impede the game. As the student completes each mission or level, the student acquires scientific skills (e.g., conversions, balancing equations, solving stoichiometry problems, etc.). In this regard, by using gamification techniques, students solve problems for themselves by carrying out "missions" that require them to apply reason, sensation, and then reflection.

LABMATTER (Jenkins 2017) was developed using *Unity 3D* which is an open-source game engine that comes with certain standard assets to help an individual engage in game development (www.unity3d.com). Programming involves using C# (i.e., C-Sharp) and JavaScript. The game engine enables the user to develop a game without having to program every single detail. The purchased standard assets used in this study can be found at the *Unity Asset* store that includes the environment, characters (https://www.mixamo.com/#/), models, and some of the game mechanics. An algorithm code (programmed by employing Lua) was used to constantly change the chemical compound and the "stem" numbers allowing the student to play the game and get new numbers each time. This feature enables the learner to get multiple practice attempts, therefore helping to reinforce the content learned in lecture.

The *LABMATTER* game was launched using HTML5 in combination with the developer-owned website using Blue-host (server provider) and WordPress (website development carrier). Hypertext Markup Language (HTML) was used in combination with plugins to operate the game. Modifying the browser cache was used to promote online play of the game. Talentlms (online lms = learning management system) (2018) enabled the user to login and to access the game by going to justmatter.talentlms.org. Mozilla Firefox (64-bit) and a high-speed Internet connection are required for desktop access.

Classroom Setting and Resources

The resources for each group are displayed in Fig. 27.2. *MasteringChemistry®* (Pearson 2018) is an online chemistry homework system available for students through the commercial publisher. (Student use of the homework system was not evaluated in this study.) The textbook chosen for this research pursued an "atoms-first" curriculum, and the lecture-based instruction was conducted at a large, southwestern R1 university. The assessments (exams 1, 2, 3, and 4) were given about every 4 weeks until the 12th week of the semester covering Chapters 1 through 10 in *Chemistry: Structure and Properties* (Tro 2017). *OpenStax* (2018) (Rice University,

	Group 1 (experimental)	**Group 2 (traditional)**	**Group 3 (control)**
Intervention	*LABMATTER (game)*	*None (opted out)*	*Online videos*
Homework	*Mastering Chemistry*	*Mastering Chemistry*	*Mastering Chemistry*
Textbook	Tro (2017)	Tro (2017)	Tro (2017)
Resource	OpenStax	OpenStax	OpenStax
Open-source App	MolView	MolView	MolView
Assessments	Instructor developed	Instructor developed	Instructor developed

Fig. 27.2 Classroom resources

peer reviewed) is an open-book source that provides students with instant access, and students also had access to the structure-drawing software, *MolView* (Bergwerf 2018), that helps with drawing Lewis structures and molecular geometry. Figure 27.3 lists the chapters covered in each assessment period with the corresponding game level and chemistry topics addressed.

Methodology

The target population consisted of students enrolled ($N = 133$) in the spring semester of first-semester general chemistry (Chem I). Verbal instructions were given to the students in the classroom setting on how to play the game as well as how to access other needed resources. The student (or player) created an account at the *Talentlms* website where they were given a list of random numbers that decided which students would play the game and which would be given access to the online videos (time-on-task equivalent control). The *LABMATTER* game and online videos were provided to the students in the first week of class where individuals subsequently signed into their specified group (Fig. 27.2): group 1 represented the game-level students, while group 3 was the corresponding online video activity group. Non-compliant students by default became the unintended group for comparison purposes subsequently referred to as the traditional group or group 2.

All students were given an open-ended, calculator-free pretest at the beginning of the semester to assess their "numerical automaticity" called the *MUST* (*Math-Up Skills Test*) (Albaladejo et al. 2018) that was scored based on the number of questions correct out of a maximum of 20 possible questions. This 15-min quiz has been

Chapter	**Assessment (Game level)**	**Topics**
Chapter 1, 2, and, 3	Exam 1 (Level 1)	Atomic structure, basic conversions, significant figures, quantum numbers, and the electromagnetic spectrum
Chapter 4 and 5	Exam 2 (Level 2)	Molar mass, moles, percentages, Lewis structures, electron configuration, molecular geometry, and electron pair geometry
Chapter 6, 7, and 8	Exam 3 (Level 3)	Redox, stoichiometry, acids and bases, and types of reactions
Chapter 9 and 10	Exam 4 (Level 4)	Thermodynamics and gas laws

Fig. 27.3 Exam content and game level

shown to be highly reliable and strongly correlated to Chem I outcomes. The *MUST* was used in a statewide study that included the present university.

Multiple-choice exam data were collected from Scantrons®. The criteria for average success are the following: *MUST* scores of 6.95/20 or 34.75% and above and course average of 70.0% and above. The assessments required students to synthesize, analyze, and make judgments using higher-order thinking skills with measurable levels. The time spent on the game and an online activity was recorded using *Talentlms*® software.

Demographic information was collected to determine the baseline characteristics of the class using *Talentlms* and to collect students' attitudes toward learning by using an e-system instrument called the *System Usability Scale (SUS)* (Brooke 1996). The *SUS* was administered after the first exam and at the end of the semester following the fourth exam to confirm the students' attitudes toward using active learning methods. All students in the class ($N = 133$) had access to the same course material, but 10 failed to complete the demographic information, and therefore a smaller population ($n = 123$) was used to describe the characteristics of the students available as reported in Table 27.1. Data from all 133 students were used to evaluate the active learning interventions. Table 27.1 displays data pertaining to the students' gender, *MUST* scores, classification, ethnicity, employment, and status as first-generation students.

Demographic data indicate that males ($n = 51$ or 41.5%) outperformed females on the *MUST* quiz given to all students at the start of the semester suggesting that males enter Chem I with greater arithmetic automaticity than females. The majority (69.1%) were classified as lower-level (freshmen + sophomores) students, and more students self-reported as white (W), non-Hispanic than the listed minorities (i.e., Hispanic (H), black (B), and other (O)). The majority of the students were not employed (61.8%), but for those who were employed, more worked off-campus than on (38/46 = 82.6%). Over 25% (31/123) reported that no parent or guardian had obtained a degree, thereby establishing their status as first-generation students.

Several analyses of demographic variables and measures of prior knowledge support each of the following conclusions: (a) lower-division males entered the course with stronger automaticity skills than corresponding females; (b) approximately the same proportion of students in each treatment group benefited from a high school chemistry course based on AP/IB, Pre-AP, and regular curricula; and (c)

Table 27.1 Demographics ($n = 123$)

			Classification				Ethnicity				Employment[a]			First-gen	
Gender	*n*	*MUST*	Fr	So	Jr	Sr	W	H	O	B	Not	On	Off	Yes	No
Male	51	8.63**	16	19	13	3	25	11	7	8	32	2	16	12	39
Female	72	5.76	28	22	15	7	25	25	11	11	44	6	22	19	53
Total	123	6.95	44	41	28	10	50	36	18	19	76	8	38	31	92

Abbreviations: *max* maximum score, *Fr* freshman, *So* sophomore, *Jr.* junior, *Sr* senior, *W* white, *H* Hispanic, *B* black, *O* other, *On/Off* employed on and off campus, *gen* generation

[a]One male did not report employment status

**$p < 0.05$ (males statistically outperformed females) those enrolled in groups 1 (game) and 3 (online videos) had significantly stronger mathematical skills than those enrolled in group 2 (traditional).

Research Questions

The semester exams were used to determine the effect and significance of the game, online activity, and the traditional treatment. The dependent variables were the four midterm exam scores, while the student group (i.e., active learning engagement vs traditional) was the independent variable.

The research questions addressed are:

Q1: What differences exist in the course averages of students who used *LABMATTER* versus those who used the online videos and those who only experienced a traditional curriculum approach without a noted High-Impact Practice (HIP)?
Q2: What differences exist in students' performance on each of the four semester exams between treatment groups?
Q3: How do students perceive the efficacy of *LABMATTER* as a tool for enhancing problem-solving in general chemistry?

Statistical Analyses

The *IBM® Statistical Packages for Social Sciences® Statistics Version 24 (SPSS®)* software was used to carry out Pearson correlation, one-way analysis of variance (ANOVA), and post hoc (Tukey) analysis. The Pearson correlation is able to identify grouping of variables that have a statistical difference. A Pearson correlation with a two-tailed test was performed at alpha level of 0.05 to compare the exam scores of the experimental group (game), the traditional group, and the control group (online activity).

For evaluation of students' perceptions (i.e., effectiveness, efficiency, and satisfaction), the *System Usability Scale (SUS)* produced a single number to represent the overall measure and usability of the system studied. A *SUS* score of 90 is interpreted as excellent, 80 is good, and anything below 70 is unacceptable (Brooke 1996).

Results

Research Question #1

Q1: What differences exist in the course averages of students who used *LABMATTER* versus those who used the online videos and those who only experienced a traditional curriculum approach without a noted High-Impact Practice (HIP)?

Tables 27.2 and 27.3 summarize the mean total scores of midterm examinations by treatment group and disaggregated by gender. Students enrolled in the game treatment (group 1) and the video treatment (group 3) outperformed those in the traditional treatment (group 2). However, no differences were found between those in the game and video treatments.

Disaggregation of the exam scores by gender revealed that both males and females in the game and video treatments outperformed males and females in the traditional treatment, but no differences were found between those enrolled in the game and video treatments.

Research Question #2

Q2: What differences exist in students' performance on each of the four semester exams between treatment groups?

Exam 1 covered Chapters 1–4 in *Atoms First*. As demonstrated in Table 27.4, there is a clear distinction between the groups. Comparisons between groups 1 and 2 and between groups 2 and 3 reveal a statistically significant difference ($p < 0.05$). Groups 1 and 3 display no statistically significant difference ($p = 0.841$). Post hoc analyses were conducted on exams 1, 2, 3, and 4, but only those from exams 1 and 2 are reported below.

A Pearson correlation of 0.179 indicated a relatively weak relationship between tests and time in the game and video treatments using two-tailed analysis. However, there was a statistically significant difference between each group investigating test versus time with an alpha level of 0.040 indicating that the more time spent engaged, the higher the test average.

Table 27.2 Average *MUST* scores for each group

	Group 1	Group 2	Group 3
n	60	18	50
MUST	7.28*	6.84	6.94*
Course Avg	73.61*	57.44	70.87*

*$p < 0.05$ (statistical differences in performance exist between Groups 1 and 2 where Group 1 outperformed Group 2 on the *MUST* and course average as did Group 3 over Group 2; Group 1 outperformed Group 3 but the difference was not statistically significant)

Table 27.3 Performance by gender ($n = 133$)

Assignment	Gender	*n* (%)	*Course Avg (*SD*)
Group 1	Female	35 (55.6)	71.7 (15.6)
	Male	28 (44.4)	76.0 (13.8)
Group 2	Female	11 (57.9)	51.0 (24.6)
	Male	8 (42.1)	66.3 (21.0)
Group 3	Female	34 (66.7)	71.2 (16.9)
	Male	17 (33.3)	70.2 (16.0)

*$p < 0.05$ (Both male and females engaged with active learning modalities outperformed students in the traditional Group 2 where students were not actively engaged in course activities)

Table 27.4 Exam 1 comparison between groups

Groups (I)	Groups (J)	Mean difference (I – J)	Std. error	Sig.	95% Confidence interval	
					Lower bound	Upper bound
1	2	19.79032*	6.57269	0.009	4.2074	35.3733
	3	2.36175	4.21105	0.841	−7.6221	12.3456
2	1	−19.79032*	6.57269	0.009	−35.3733	−4.2074
	3	−17.42857*	6.64093	0.026	−33.1733	−1.6839
3	1	−2.36175	4.21105	0.841	−12.3456	7.6221
	2	17.42857*	6.64093	0.026	1.6839	33.1733

*Mean difference is significant at the $p < 0.05$ level

Table 27.5 Exam 2 comparison between groups

Groups (I)	Groups (J)	Mean difference (I – J)	Std. error	Sig.	95% Confidence interval	
					Lower bound	Upper bound
1	2	20.53226*	6.18437	0.003	5.8700	35.1945
	3	0.28226	3.96226	0.997	−9.1117	9.6762
2	1	−20.53226*	6.18437	0.003	−35.1945	−5.8700
	3	−20.25000*	6.24858	0.004	−35.0645	−5.4355
3	1	−0.28226	3.96226	0.997	−9.6762	9.1117
	2	20.25000*	6.24858	0.004	5.4355	35.0645

*$p < 0.05$

On the second exam, covering Chapters 4–5 in *Atoms First*, there was also a statistically significant difference ($p < 0.05$) between student groups. A post hoc (Tukey) analysis was conducted on the study's three groups yielding a statistically significant difference of 0.003 (Table 27.5). Groups 1 and 3 demonstrated no statistically significant difference resulting in a of value of 0.997. However, as with the first midterm exam, comparisons of groups 1 and 2 and groups 2 and 3 revealed statistically significant differences.

Pearson correlation was conducted using a test versus time analysis. A quantitative value of 0.288 was found between test and time (in the game and video treatments). This result implies that there is a highly positive correlation between study time and achievement scores. The significance of the relationship between study time and achievement within groups 1 and 3 are promising indicators that positively affect students.

The third exam covered Chapters 6–8 in *Atoms First*, and the fourth exam covered the remaining two chapters, 9–10. Like the first two exam units, there was a statistically significant difference ($p < 0.05$) between the active learning groups (1 and 3) and group 2. However, the Pearson correlations for exams 3 and 4 produced no statistically significant relationships ($p < 0.05$) and very low positive correlations (0.093 and 0.025) between the exams and the time engaged.

Research Question #3

Q3: How do students perceive the efficacy of *LABMATTER* as a tool for enhancing problem-solving in general chemistry?

Research question 3 analyzed student's attitude toward *LABMATTER* using the *SUS*. There are ten items on the *SUS* to which the students responded using a Likert positive/negative scale to produce a score revealing how students felt about using the game. The students perceived the game as a useful medium for learning chemistry with an average *SUS* score of 70.3 out of a maximum score of 100. This score was deemed acceptable according to published suggestions (Brooke 1996). As the game becomes more sophisticated and programming glitches improve, so will students' attitudes about playing this game.

Course Outcomes

As a general addendum to this study, Table 27.6 reports grades for successful students (grades of A, B, or C) and those deemed unsuccessful (grades of D, F, or W). The overall grades for group 1 compared favorably to students in group 3 (63.9% vs 57.7% successful) and both groups outperformed students in group 2 (33.3% successful).

Summary

The results of this study suggest that *LABMATTER* offers students an opportunity to practice problem-solving skills relevant to introductory college chemistry in a more inviting medium, thereby, if only slightly, increasing students' chances of success over online academic videos in widespread use and published by commercial sources. Interactive games for students can provide stimulating, rewarding, and even exciting venues that significantly improve retention of scientific material, maximize engagement, and improve learning. Wieman (2014) observed that "active learning methods achieve better education outcomes" (p. 8319). When students are engaged and motivated, classroom performance improves, and active learning is one of the best ways to accomplish meaningful engagement.

Table 27.6 Course grades by group

	Group 1	Group 2	Group 3
Course grade	*n* (average)	*n* (average)	*n* (average)
A	6 (16.7%)	0	3 (11.5%)
B	10 (27.8%)	2 (22.2%)	9 (34.6%)
C	7 (19.4%)	1 (11.1%)	3 (11.5%)
Successful	**23 (63.9%)**	**3 (33.3%)**	**15 (57.7%)**
D	7 (19.4%)	2 (22.2%)	5 (19.2%)
F	5 (13.9%)	4 (44.4%)	6 (23.1%)
W	1 (2.8%)	0	0
Unsuccessful	**13 (36.1%)**	**6 (66.7%)**	**11 (42.3%)**
Total	**36**	**9**	**26**

References

Albaladejo, J. D. P., Broadway, S., Mamiya, B., Petros, A., Powell, C. B., Shelton, G. R., Walker, D. R., Weber, R., Williamson, V. M., & Mason, D. (2018). ConfChem conference on mathematics in undergraduate chemistry instruction: MUST-know pilot study—Math preparation study from Texas. *Journal of Chemical Education, 95*(8), 1428–1429. https://doi.org/10.1021/acs.jchemed.8b00096 [Articles ASAP (As Soon As Publishable): July 20, 2018 (Report).].

Bergwerf, H. (2018). *MolView*. Retrieved from https://Molview.org. Accessed 8 Nov 2018.

Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability Evaluation in Industry, 189(194), 4–7. Retrieved from https://hell.meiert.org/core/pdf/sus.pdf. Accessed 17 Nov 2019.

Civic Enterprises in association with Peter D. Hart Research Associates. (2006, March). *The silent epidemic: Perspectives of high school dropouts*. Seattle: The Bill & Melinda Gates Foundation. Retrieved from https://www.google.com/search?client=safari&rls=en&q=www.gatesfoundation.org/united-states+/Documents/TheSilentEpidemic3-06FINAL.pdf&ie=UTF-8&oe=UTF-8. Accessed 8 Nov 2018.

Clark, K., & Sheridan, K. (2010). Game design through mentoring and collaboration. *Journal of Educational Multimedia and Hypermedia, 19*(2), 5–22.

Clark, K., Brandt, J., Hopkins, R., & Wilhelm, J. (2010). Making games after school: Participatory game design in non-formal learning environments. *Educational Technology, 49*(6), 40–44.

ESA (Entertainment Software Association). (2012). *2012 Sales, demographic and usage data: Essential facts about the computer and video game industry*. Retrieved from http://www.theesa.com/article/essential-facts-computer-video-game-industry-2018/. Accessed 5 Nov 2018.

Gutnick, A. L., Robb, M., Takeuchi, L., & Kotler, J. (2011). *Always connected: The new digital media habits of young children*. New York: Joan Ganz Cooney Center at Sesame Workshop. Retrieved from www.joanganzcooneycenter.org/publication /always-connected-the-new-digital-media-habits-of-young-children/. Accessed 8 Nov 2018.

Jenkins, D. (2017). LABMATTERä. Retrieved from https://justmatter.org. Accessed 8 Nov 2018.

OpenStax. (2018). Rice University. Retrieved from https://openstax.org/details/books/chemistry-atoms-first. Accessed 8 Nov 2018.

Pearson. (2018). Mastering chemistry. *Pearsonmylabandmastering*. Retrieved from https://www.pearsonmylabandmastering.com/northamerica/masteringchemistry. Accessed 8 Nov 2018.

Rideout, V. J. (2011). *Zero to eight: Children's media use in America*. San Francisco: Common Sense Media.

Rideout, V. J., Foehr, U. G., & Roberts, D. F. (2010). *Generation M2: Media in the lives of 8- to 18-year- olds*. Menlo Park: Kaiser Family Foundation.

Shell, D. F., Brooks, D. W., Trainin, G., Wilson, K. M., Kauffman, D. F., & Herr, L. M. (2010). *The unified learning model* (pp. 1–4). Dordrecht: Springer.

Talentlms. https://www.talentlms.com. Accessed 8 Nov 2018.

Toprac, P. (2018). Six predictions about the future of gaming from a computer scientist. Interview for *The Texas Scientist*. College of Natural Science, The University of Texas at Austin. https://cns.utexas.edu/news/six-predictions-about-the-future-of-gaming-from-a-computer-scientist?utm_campaign=NASC_FY18-19_Newsletter_TexasScience-Oct2018_EML&utm_medium=email&utm_source=Eloqua. Accessed 5 Nov 2018.

Tro, N. J. (2017). *Chemistry: Structure and properties* (2nd ed.). Hoboken: Pearson Education, Inc.

Wieman, C. E. (2014). Large-scale comparison of science teaching methods sends clear message. Commentary. *Proceedings of the National Academy of Sciences of the United States of America, 111*(23), 8319–8320.

Williamson, V. M. Pages 67–78 in Bunce D. and Cole R., (ed.), (2008). *Nuts and bolts of chemical education research*. Washington, DC: American Chemical Society.