

Chemistry Education Research and Practice

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Learning beyond the classroom: Using text messages to measure General Chemistry students' study habits

Lewis, Scott*; Ye, Li; Oueini, Razanne; Dickerson, Austin

University of South Florida - Chemistry, Tampa, Florida, United States

* 4202 E. Fowler Avenue CHE205 , Tampa, Florida 33620

United States.

T: 813-974-3099 F: 813-974-3203. email: slewis@usf.edu

Introduction

Understanding factors related to student learning in General Chemistry is necessary to design and evaluate implementations to improve academic performance. Considerable effort has been made toward this end through the use of reformed pedagogical techniques. These techniques target in-class activities and have shown a notable impact on metrics for academic performance (Freeman *et al.*, 2014). In post-secondary education however, students spend only three to five hours per week in class with the opportunity to spend considerably more time outside of class studying the course materials. This leads to two overlapping possibilities regarding a causal explanation for the effectiveness of in-class pedagogical reform: i) students' experiences in class cause learning gains or ii) the reform modifies students' activities outside of class that cause learning gains. Currently little information is known regarding post-secondary chemistry students' studying of course material outside of class, herein referred to as study habits. This study seeks to examine a novel method for measuring students' study habits and explore the role of study habits in academic performance.

Background

Past Work on Study Habits

Considerable past work on college students' study habits has been carried out in the fields of education and psychology. In a recent meta-analysis of the work on study habits, Crede and Kuncel (2008) described the empirical and theoretical literature on studying behaviors as fragmented. They organized studying behaviors based on the constructs: *study skills*, knowing how to study, *study habits*, the frequency and type of actions taken toward studying and *study attitudes*, the motivation toward studying. These constructs are differentiated from study processes that describe the depth of processing on a continuum from deep (an effort to relate new material to previously learned contexts) to surface (characterized as memorization without seeking context). In the meta-analysis, the researchers

1
2
3 identified 40 studies relating study habits to college GPA and found correlations that
4 average approximately 0.33 with a 90% interval between 0.09 and 0.51.

5 Relationships between study habits and individual course performance was lower,
6 averaging 0.26, which the authors attribute to not being able to correct for reliability
7 in individual course grades. Study habits also featured a weak relationship with
8 established measures of general cognitive ability such as high school GPA or college
9 admissions tests. This suggests that the relationship between study habits and
10 academic college performance is unique from the well-established relationships
11 between measures of cognitive ability and student performance. Further, it helps to
12 rule out the explanation that stronger students exhibit better study habits and that
13 this is responsible for the observed correlation; instead it suggests that students can
14 benefit from effective study habits regardless of incoming ability.
15
16
17

18 There has been little research attention toward measuring study habits in the
19 context of post-secondary chemistry. Richards-Babb and Jackson (2011)
20 investigated gender differences in study habits through a survey given at the end of
21 General Chemistry and reported that male students were more likely to
22 procrastinate. Also related, Li *et al.* (2013) investigated post-secondary chemistry
23 students' conceptions of learning chemistry and approaches to learning chemistry.
24 Conceptions of learning chemistry were measured by a survey developed based on
25 earlier, more general research interviewing students about learning experiences.
26 This work identified memorizing, testing, calculating / practicing and higher order
27 thinking labeled as transforming as the relevant themes in students' conceptions of
28 learning. Approaches to learning were measured based on the previously described
29 study processes and the continuum from surface to deep learning. The study found
30 that students who were characterized as deep learners conceived of learning in
31 multiple ways including transforming, memorizing and testing, while learners that
32 used surface strategy employ memorizing and testing.
33
34
35

36 Most studies that investigated students' study habits used a single-admission
37 survey that may be problematic for two reasons. First, as a single measure, it
38 presumes that students' study habits are constant, whereas it is possible that
39 students' study habits adapt to the nature of the content and with familiarity toward
40 assessment expectations. Second, as an in-class survey, it relies upon retrospection
41 on behalf of the student, particularly when it is given at the end of the semester.
42 Past research has called into question the accuracy of retrospective accounts,
43 particularly at lengthier time intervals (Bernard *et al.*, 1984). By exploring study
44 habits at multiple time points, both problems may be minimized as changes over
45 time can be documented and participants would not be asked to reflect upon several
46 months of study habits.
47
48

49 There have also been efforts made to improve students' study habits. Cook *et*
50 *al.* (2013) implemented a one-day lecture for General Chemistry students that
51 presented differences in expectations between post-secondary and secondary
52 education as described by Bloom's taxonomy. They also presented metacognitive
53 learning tools including a study cycle. At the conclusion, students made a brief
54 written statement committing to use some of the tools presented. Student
55 attendance to the lecture was voluntary though students received a bonus
56 equivalent to 0.5% of their grade for attending. Attending students were compared
57
58
59
60

1
2
3 to non-attendees on the metric of points earned in the course post-implementation
4 (transformed to follow a normal distribution). The statistical comparison used
5 students' first exam score as a covariate as it preceded the implementation. The
6 results of an ANCOVA showed that students who attended the treatment performed
7 better on the outcome measure than those who did not. These results indicate the
8 potential importance of student study habits to student learning, but the results
9 could also be indicative of a confounding variable such as student motivation. For
10 example, the authors did note that the control group missed more exams than the
11 treatment group and that could be an indication that the groups differed in their
12 motivation to succeed in the class. Incorporating a measure of study habits before
13 and after the intervention would further elucidate the impact of the intervention
14 and help better establish a causal connection such as the intervention impacted
15 student study habits which led to greater student learning.
16
17

18
19 In summary, considerable research has shown a relationship between study
20 habits and academic performance in post-secondary education but not in the
21 sciences. The following study seeks to address this gap in the research literature by
22 investigating the frequency and types of study habits in General Chemistry. The
23 creation of a detailed measure of student study habits as described below can open
24 two potentially fruitful areas of study. First, measuring study habits can inform
25 efforts to better understand the factors related to academic success (study habits as
26 an independent variable). Second, measuring study habits can aid explorations of
27 instructional efforts to improve students' study habits (study habits as a dependent
28 variable). The following study takes an exploratory approach to examine study
29 habits as an independent variable.
30
31
32

33 *Experience Sampling Method*

34 The methodology used in this study is Experience Sampling Method
35 (ESM) which uses technology to measure participants' self-report of their actions or
36 psychological state while the participant is in their natural environment. What
37 follows is a brief summary of the methodology as it applies to the current study. For
38 a complete introduction to Experience Sampling Method (ESM) including
39 methodological stance, research antecedents and examples, readers are advised to
40 please see Hektner et al. (2007). By measuring in the participant's natural
41 environment, researchers can learn about participants' actions outside of a
42 particular research setting (e.g. the classroom) while relying on a much more
43 proximal retrospective account than traditionally done. ESM has been described as
44 systemic phenomenography in that the information collected relies on self-report
45 and remains restricted to describing the participants' perspective on the area of
46 focus. It is considered systemic in that ESM uses technology to facilitate multiple
47 measures of a construct from each participant to establish reliability and investigate
48 patterns within a participant.
49
50
51
52

53 ESM has been used in a variety of contexts, particularly in the field of
54 psychology where it has been used to explore constructs as diverse as morality,
55 mental illness and substance abuse.(Hoffman et al., 2014; Smyth and Stone, 2003) It
56 has also been used with medical applications to investigate disorders, drug abuse
57 and treatment effectiveness.(Hektner *et al.*, 2007) In education it has been used
58
59
60

1
2
3 most often at the secondary level to investigate student motivation, satisfaction with
4 the educational environment or the nature of the environment.(Csikszentmihai,
5 2014) To date, we could not locate a study that has used ESM to explore a post-
6 secondary chemistry setting or post-secondary student study habits.
7
8

9 10 **Research Questions**

11
12 As ESM has not been previously used to explore post-secondary study habits,
13 the first research goal was to establish the utility of this method to measure post-
14 secondary students' study habits. Additional research goals include relating study
15 habits to academic performance, which speaks to the relative importance of study
16 habits, and investigating the extent study habits change. If study habits are found to
17 change over the semester, this would suggest a fruitful line of research to
18 investigate instructional actions to direct student study habits toward effective
19 practices. Specifically, this study was guided by the following research questions:
20

- 21 1a. To what extent is it feasible to measure student study habits using ESM?
- 22 1b. To what extent is there evidence for the validity of the data collected on
- 23 student study habits?
- 24
- 25 2. Which study habits were related to academic performance in the course as
- 26 measured by a cumulative final exam?
- 27
- 28 3. How did student study habits change over the course of the semester?
- 29

30 31 **Research Setting**

32
33 This study was conducted over one semester at a large research-intensive
34 university in the southeast United States. At the setting four classes of General
35 Chemistry were offered with class sizes between 200 and 225 students. The classes
36 are coordinated where the instructors agree to a common syllabus, textbook,
37 grading scheme, content sequence and pace. The classes also employed common
38 exams where students across all classes take the same exam at the same time. The
39 exams were constructed by contributions from each of the four instructors and used
40 multiple-choice questions and a measure of linked concepts (Ye and Lewis, 2014).
41 The measure of linked concepts provides a brief description of a chemical situation
42 and has students evaluate the legitimacy of a series of statements as true or false.
43 The series of statements span the content throughout the course and are meant to
44 have students consider how concepts throughout the course are linked.
45
46

47 To aid student studying, past exams were posted approximately two weeks
48 before the actual exam and are referred to as practice tests. The textbook used was
49 Tro's *Chemistry: A Molecular Approach* (2014) and the content sequence was:
50 quantum numbers, periodic trends, Lewis structures, shapes and polarity, gas laws,
51 thermodynamics, intermolecular forces and properties of solutions. Grades were
52 determined largely by performance on three in-class exams (15% each) and the
53 cumulative final exam (25%). A smaller portion of the grade was attributed to
54 three different effort-based measures at 10% each. First, the class used weekly
55 peer-led problem-solving sessions where students worked in groups on problems
56 designed by the instructors with the aid of peer-leaders (Lewis, 2011). Attendance
57
58
59
60

1
2
3 and participation in these sessions was worth 10% of their grade. Second, students
4 were graded on their performance on eight online homework assignments using
5 Sapling Learning. Third, instructors used clickers to facilitate in-class questions in
6 the large lecture-hall setting.
7
8

9 **Methods**

10
11
12 Students were recruited for this study from three of the four General
13 Chemistry classes at the research setting. One class was omitted from this study as
14 the instructor for the class was a member of the research team, and there was
15 concern that recruitment might appear coercive to students. Among the three
16 classes that were recruited, 670 students were enrolled (out of 889 students among
17 the four classes). Recruitment occurred on the first day of class by describing the
18 nature of participation in this study. Participants would be asked for their cell
19 phone number and would periodically receive a text message that inquired “Have
20 you studied for General Chemistry I in the past 48 hours? If so, how did you study?”
21 The text messages would be sent approximately twice a week at random times
22 between 9 AM and 9 PM. Participants would be asked to respond to the message
23 within 12 hours of receipt if possible and were given an instruction sheet that
24 included example responses. To encourage participation, students who responded
25 to 80% or more of the text messages would be entered into a raffle for a \$25 gift
26 card at the end of the semester. The university’s Institutional Review Board
27 approved these procedures. The recruitment effort led to 301 participants
28 consenting of the 670 students (44.9%). The text message inquiry was sent out as
29 described above 28 times over the course of the semester.
30
31
32
33

34 Student responses to the text messages were combined with data collected in
35 the normal educational setting from either university records or in-class records.
36 This data includes student responses to the revised two-factor Study Process
37 Questionnaire (rSPQ) administered on the first day of class (Biggs, 2001). The rSPQ
38 is a 20 item Likert-scale instrument meant to measure students’ study processes.
39 The instrument was revised by the original instrument’s author and measures
40 respondents on two sub-scales: deep approach and surface approach. The deep
41 approach can be characterized by intrinsic interest or a motivation to understand.
42 Example items from the rSPQ that measure the deep approach are “I come to most
43 classes with questions in mind that I want answering” and “I find that at time
44 studying gives me a feeling of deep personal satisfaction.” The surface approach can
45 be characterized by a narrow focus on content and memorization with example
46 statements “My aim is to pass the course while doing as little work as possible” and
47 “I find the best way to pass examinations is to try to remember answers to likely
48 questions.” In this study, students were asked to consider their study habits in
49 general, but if they need to consider a subject, consider how they would study for
50 chemistry or a science course. The Likert-scale was a five-point range from “this
51 item is never or only rarely true of me” to “this item is always or almost always true
52 of me,” and each factor score represents the sum of ten associated items.
53
54
55
56

57 In this study the rSPQ is thought to measure the quality of studying where
58 the deep approach describes the desirable educational process (Biggs 2001). This is
59
60

differentiated from study habits, which describe the type and frequency of studying. There are expected relationships between the constructs, for example students who employ a deep approach are expected to study more frequently. Student scores on the rSPQ are considered in contrast to their cohort as recommended by Biggs (2001). Additional measures include student demographics and SAT scores (a measure of incoming college preparation) from university records, and student performance on exams from in-class records. Descriptive statistics on each of these measures for the population of 899 students are presented in Table 1.

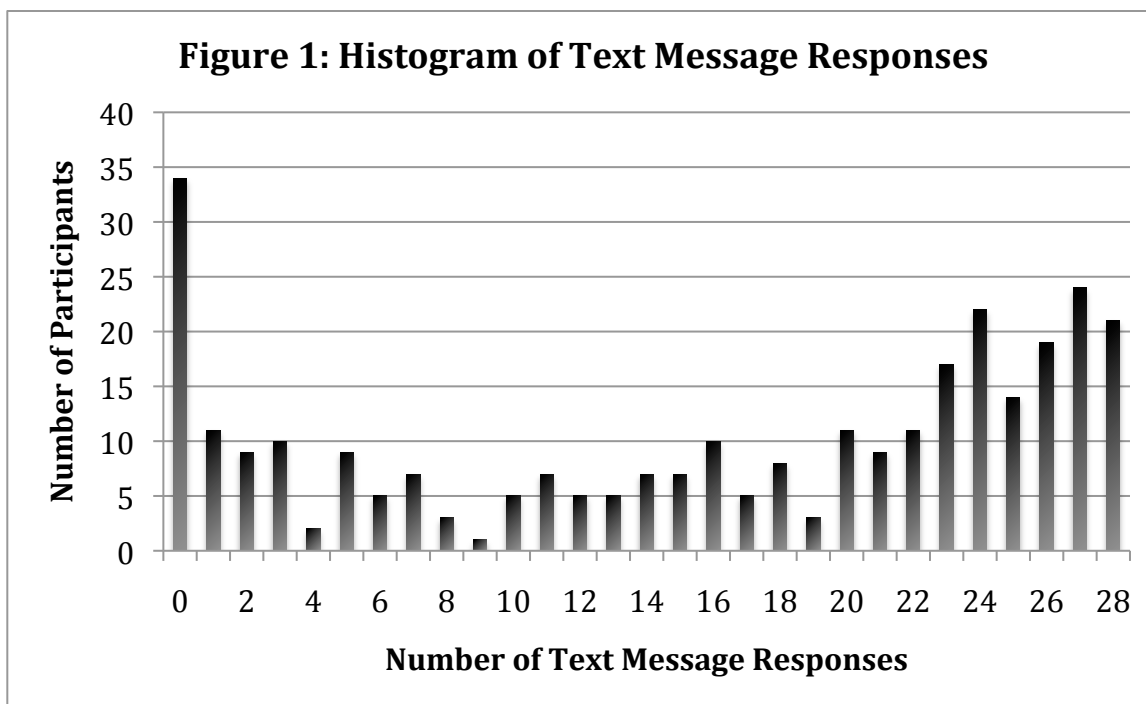
Table 1: Descriptive Statistics of Measures

Variable	Average	St. Dev.	N	Theoretical Range	Cronbach's α
SAT Math	552	68	695	[200, 800]	
SAT Verbal	548	73	695	[200, 800]	
Deep Approach	32.1	6.7	797	[10, 50]	0.826
Surface Approach	23.8	6.2	797	[10, 50]	0.776
Final Exam	48.5	15.7	754	[0, 100]	0.816

Results and Discussion

Feasibility of ESM for Measuring Study Habits

Over the course of the semester 4,775 responses were collected in response to the 28 inquiries. This represents an average of approximately 16 responses per participant. A histogram of the frequency of student responses is presented in Figure 1. From the histogram, there were 34 students (11.3% of participants) who never responded to the text message inquiries. There were also a sizable number of students who regularly responded, as 188 students (62.5%) responded to at least half of the messages and 137 students (45.5%) responded to at least three-quarters of the messages.



In terms of feasibility, it is plausible to have a substantial portion of the recruited population respond to this data collection technique. It is worth noting that the raffle incentive required responses to 80% of the inquiries (23 or more inquiries), which may partially explain the rise in number of participants who responded to 23 or more inquiries. Instead, the largest source of data attrition in the study was during the initial recruitment, where out of an initial population of 670 students, 301 students agreed to participate (44.9%). This suggests that future research studies that intend to rely on a large number of responses would benefit by planning for a substantially larger recruitment pool. The current data indicates that roughly one-quarter of the initially recruited population provided responses to at least half of the inquiries.

At the close of the semester, an additional text message inquiry was sent asking participants if they would participate in a similar study using text messages in the future. Of the 94 respondents, 78% responded positively compared to 18% negative (with the remainder unsure). The most common negative comment (5 responses) was that students found the messages annoying. However, the most common comment (52 responses) described the convenience in participating with some indicating it was less of a time commitment compared to traditional studies.

Evidence for Validity

First, to determine how generalizable the sample is, participants were compared to the non-participants on each variable describing an in-coming characteristic: SAT sub-scores and the Surface Approach and Deep Approach score from the rSPQ. Scores on each metric are compared in Table 2. Using the two one-sided t-tests method (Lewis and Lewis, 2008) for establishing equivalence, with an equivalence interval equal to the small effect size (Cohen's $d = 0.2$), the two groups

were equivalent on Math SAT and the Surface Approach. (Cohen, 1988) The departures from equivalence were minor and when the interval was expanded to $d = 0.25$ the two groups were equivalent on all metrics. Participants were also compared to non-participants based on demographic characteristics of gender and minority status. For this comparison the chi-square test was used with the effect size estimated using Cohen's w . The comparison found that both differences were less than a small effect, which Cohen operationalized as $w = 0.10$. For gender $\chi^2 = 3.27$, $w = 0.06$, and for minority status $\chi^2 = 2.39$, $w = 0.05$. The above comparisons serve to investigate self-selection bias in this study and found only small departures from the participants and the non-participants on the measures considered. These measures only serve as an indirect measure of self-selection bias as it is still possible that the study habits of the participants and non-participants differ and study habits of non-participants could not be investigated with the data collected. It is therefore proposed that no evidence was found to believe that the sample is biased by self-selection and the sample may be generalizable to the population of General Chemistry students at the setting.¹

Table 2: Generalizable Validity: Comparing Participants to Non-participants

	Participants		Non-participants
	Average (St. Dev.) <i>N</i>		Average (St. Dev.) <i>N</i>
SAT Math	554 (65) 237		550 (69) 458
SAT Verbal	552 (70) 237		545 (74) 458
Deep Approach	32.7 (6.6) 288		31.8 (6.7) 509
Surface Approach	23.6 (6.3) 288		23.9 (6.2) 509
	Participant Demographics (<i>N</i>)		Non-Participant Demographics (<i>N</i>)
Gender	65% Female (300)		59% Female (594)
Minority	45% URM* (277)		40% URM* (559)

*URM = under-represented minority (as defined by the National Science Foundation)

To explore the content of the responses, the text messages were coded using an open-coding scheme. The coding process resulted in 16 codes as shown in Table 3. Each response was coded and responses could be coded with multiple codes. For example, "Yes the back of book problems, reading the chapter, and doing the online homework assignment" was coded for Textbook, Practice Problems and Homework.

¹ Analyses presented later in this manuscript will rely on a subset of the sample based on frequencies of responses to text messages. The correlation between frequency of responses and each of the measures in Table 2 were found to be weak, with $r = 0.16$ for response rate to Math SAT or Verbal SAT and $|r| < 0.14$ for the other measures, indicating subsets generally continue with minor departures from the population.

To check the inter-rater reliability of the coding scheme, 10% of the text messages were randomly selected and coded by a researcher who was independent of the first coding pass. The resulting codes agreed with the original code for 94% of the responses. Table 3 also presents the relative frequency of the codes as the percent of responses that used a particular code.

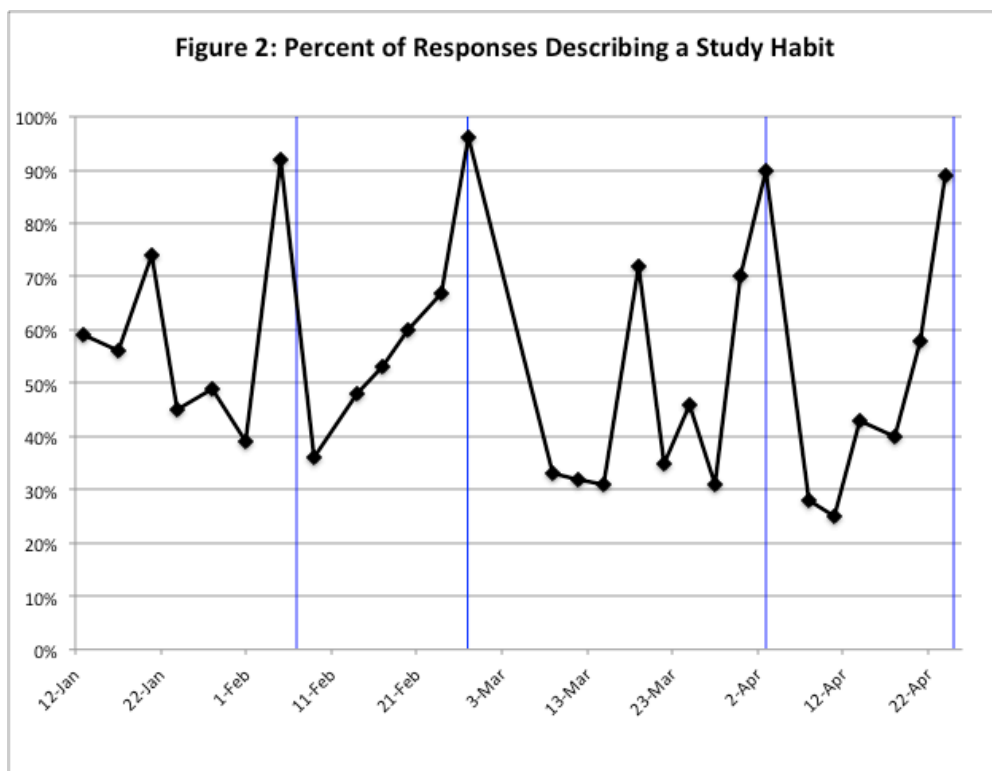
Table 3: Types of Study Habits and Frequency

Code	Percent
Did not study	42.2%
Reviewed notes or PowerPoint	18.8%
Reviewed the textbook	16.4%
Online homework	14.2%
Practiced problems	6.8%
Previous exams or study guides	5.7%
Unspecified yes	4.5%
Used online materials	2.6%
Worked with friends or in a group	2.4%
Attended peer leading or reviewed peer leading assignment	2.1%
Worked with a tutor	1.9%
Attended class	1.1%
Made flashcards	0.9%
Visited professor	0.3%
Attended lab	0.2%
Reviewed tables, models or charts	0.2%

Text messages were then also coded dichotomously as either a study habit was used or the participant did not study. The codes unspecified yes, attended class, attended peer leading or attended lab were coded as missing in this categorization as it was not clear whether these participants had employed a study habit. With the new dichotomous codes, the percent of participants employing a study approach was determined for each text message inquiry. The percent of participants using a study habit is plotted by date in Figure 2, using only those participants who responded to at least half of the 28 inquiries ($N = 188$). The vertical lines in Figure 2 correspond to the exam dates in the class. From Figure 2, the percent of students who report studying increases leading up to each exam, peaks at the exam date and subsequently drops-off. This matches the expected pattern of instructional experience where student inquiries tend to ramp up leading up to the date of an exam, lending content validity to the responses received.

Ultimately, the measure of students' study habits proposed is still reliant on self-report. Self-reported data may be influenced by factors that cannot be ruled out such as participants' belief in a socially desirable response pattern or errors in participants' efforts to recollect. Such factors would impact the accuracy of the responses as a measure of actual student actions. As study habits by definition occur outside of a controlled research setting, attempts to triangulate the measure without relying on self-report would require extensive observations that would

impose on participants' privacy. This serves as an unavoidable limitation of this study though it is proposed that participants' self-report of study habits do offer value in understanding the factors needed for successful academic performance.



Relationship of Study Habits to Academic Performance

Identifying successful study habits can guide efforts to improve study habits through student advising. The knowledge of productive study habits can also inform evaluations of reform pedagogies allowing an exploration of the extent reform pedagogies promote effective study habits. To investigate the relationship of study habits to academic performance, each participant was characterized by the percent of the participant's responses that indicated each of the study habits shown in Table 3. Second, only participants who responded to half of the text message inquiries were considered to promote stability in the percentage. That is, a participant who indicated reviewing notes in 14 out of 21 responses indicates a more stable pattern than another participant who indicated so in 2 out of 3 responses. Finally, academic performance was operationalized by performance on the cumulative final exam discussed earlier. This measure was chosen as the clusters represent study habits across the semester and the final exam was the only measure to occur at the end of the semester.

Initially, correlations between each study habit and the final exam were conducted. Each correlation indicated a weak relationship with the strongest relationship of 0.14 between percent of responses using the textbook and final exam score. Since correlations only indicate the strength of a linear relationship, the data

was further explored for the possibility of relationships that do not follow a linear pattern. Owing to the substantial number of study habits present, the decision was made to conduct a cluster analysis to look for patterns among the multiple study habits. Cluster analysis is an algorithm that measures the distance between each case (student) on the variables (frequency of study habits) and combines pairs of students who feature the smallest distance into a cluster. The algorithm continues to combine students and clusters of students until it reaches a user-specified number of clusters. In this way, cluster analysis can be used to find groups of students who have similar profiles across multiple variables (Everitt *et al.*, 2011). Cluster analysis can be used to describe the data in terms of number of students per group and the average study habits within each group. These groupings can then facilitate investigating relationships among other measures.

For the cluster analysis, only the six most prevalent study habits in Table 3 were used, as these were each represented by at least 5% and were also readily interpretable (the next most prevalent code would be the unspecified yes). A hierarchical cluster analysis using Ward's method and squared Euclidean distance was employed to create clusters that were distinct from each other (Aldenderfer and Blashfield, 1984). To determine the number of clusters, the cluster analysis began with six clusters that were evaluated based on sample size in each cluster and the average percent for each study habit. Then an analysis to create five clusters was conducted to determine which two clusters were combined; these clusters were evaluated based on qualitative similarity on study habit percentages and the relative sample size of each cluster. The analysis was continued until reducing the number of clusters meant losing a cluster that was substantially distinct. The intent was to determine the number of clusters that led to reasonable representation in the sample for each cluster and where each cluster was distinct. This resulted in three clusters that are characterized in Table 4.

Table 4: Cluster Analysis Results – Study Habits

	Cluster 1	Cluster 2	Cluster 3
Sample Size	64	62	62
Study Habits	Average (St Dev)		
Did not study	67% (13%)	<u>26% (14%)</u>	33% (15%)
Reviewed notes or PowerPoint	<u>8% (9%)</u>	22% (22%)	26% (16%)
Reviewed the textbook	<u>8% (11%)</u>	35% (16%)	<u>8% (8%)</u>
Online homework	<u>5% (7%)</u>	11% (9%)	25% (16%)
Practiced problems	4% (6%)	10% (12%)	6% (9%)
Practice tests or study guides	3% (4%)	6% (7%)	9% (8%)

Bold indicates study habit has more than +0.5 standard deviation different than the overall average; *italic underline* indicates study habit is less than -0.5 standard deviations different than the overall average

Table 4 describes three distinct clusters that arose from the study habits in the sample. To place the values in context, 14% would indicate that they used the study habit at least twice and at most four times over the course of the semester. The sample distribution among the three clusters is relatively even which suggests that each cluster has prevalence among the sample. Participants in Cluster 1 indicated not studying far more often than the rest of the sub-sample (67% versus 42% for the sub-sample) and subsequently indicated reviewing notes and the online homework less often than the sub-sample. Cluster 2 was more than one standard deviation greater than sub-sample on use of the textbook (35% versus 16%). Cluster 2 was also higher on practicing problems and lower on the percentage of not studying. Cluster 3 was noteworthy for describing the online homework as their study habit, but was also higher on reviewing notes and the practice tests.

The three clusters were compared on the five other measures with data presented in Table 5. To compare the clusters an Analysis of Variance (ANOVA) was performed with $\alpha = 0.05$ which provides a group-wise error rate of 0.23 across the five tests. The effect size was also characterized by Cohen's f , where 0.10 is a small effect and 0.25 is a medium effect (Cohen, 1988). Interestingly, neither SAT sub-score was found to be statistically significant with negligible effects for Verbal SAT ($F = 0.233$; $p = 0.792$; $f = 0.06$) and Math SAT ($F = 0.135$; $p = 0.874$; $f = 0.04$). For the study approaches, the clusters differed with medium effects on both the deep approach ($F = 4.190$; $p = 0.017$; $f = 0.22$) and the surface approach ($F = 7.315$; $p < 0.001$; $f = 0.27$). Post-hoc comparisons using the Tukey test indicate that the significant difference is Cluster 1 is higher on the surface approach than the other two clusters and Cluster 2 is higher on the deep approach than Cluster 1. On the final exam metric, the clusters were also different with a near medium effect ($F = 3.663$; $p = 0.028$; $f = 0.21$). Post-hoc analysis describes the significant difference as Cluster 2 higher than Cluster 1. An ANCOVA analysis controlling for SAT sub-scores on the final exam measure indicated similar results ($F = 3.881$; $f = 0.24$) as the original ANOVA analysis.

Table 5: Study Habit Clusters Compared

	Cluster 1	Cluster 2	Cluster 3
Variables	Average (St Dev)		
SAT Math	566 (62)	561 (71)	560 (57)
SAT Verbal	557 (70)	566 (81)	558 (71)
Deep Approach	30.4 (6.5)	33.7 (6.4)	31.8 (5.8)
Surface Approach	25.7 (5.9)	22.1 (6.1)	22.4 (5.6)
Final Exam	43.8 (13.3)	51.1 (15.2)	46.8 (14.0)

1
2
3
4 Thus, it appears that Cluster 2, which comprises roughly one-third of the
5 sample, had higher scores on average on the final exam measure. This suggests the
6 study approaches described as reviewing the textbook and practicing problems
7 leads to increased academic performance in the course. Not surprisingly, Cluster 1,
8 which indicated predominately not studying, performed worse. That Cluster 2
9 scored higher on the deep approach and Cluster 1 on the surface approach lends
10 external validity to the qualitative difference between these two groups. Cluster 3's
11 performance on the final exam is interesting as it was comparable to Cluster 1. The
12 study efforts of Cluster 3 are more concentrated on the online homework. It is
13 hypothesized that these students *perceived* the completion of the required online
14 homework as suitable preparation for the exams. The central feature of the
15 hypothesis is the emphasis on perception. Since the online homework was a
16 mandatory part of the class it is likely that the strong majority of students
17 completed it, however the students in Cluster 3 may have perceived it as
18 satisfactory preparation whereas students in Cluster 2 believed that additional
19 preparation was necessary. Thus, Cluster 1 may be described as knowingly not
20 studying, Cluster 3 as believing the required course components constitute
21 satisfactory studying and Cluster 2 studied in addition to the required course
22 components by relying on the textbook and practicing problems.
23
24

25
26 The relationship of the study habits measured by ESM to a measure of
27 academic performance serves as support for external validity of the data collected.
28 The finding that students who study more regularly perform better on the
29 cumulative final exam may not be surprising. However, the findings that
30 approximately one-third of the sample study regularly, which matches the baseline
31 observed in Figure 2, is of importance as it suggests that there is ample ground for
32 promoting effective study habits. That the students who study regularly are also not
33 distinguishable from the other groups based on SAT scores also partially rules out
34 the competing explanation that these students were more academically prepared
35 prior to the semester. Another possible explanation for differences in academic
36 performance may include differences between clusters in student motivation to
37 succeed in the course; in particular, it is plausible that differences in motivation may
38 manifest themselves in more frequent studying.
39
40
41
42

43 *Study Habits Change over the Semester*

44

45
46 To investigate changes in study habits over the course of the semester, the
47 analysis focused on the text message inquiries that were sent out immediately
48 preceding each exam. The decision to focus on these four text message inquiries
49 was based on the increase observed in describable study habits that coincided with
50 the exams as shown in Figure 2. It will also lend the most insight into students'
51 exam preparation strategies. As a measure of change in study habits over the
52 semester, the analysis was conducted on only the 113 participants who responded
53 to each of the four messages in question; otherwise, observed changes could result
54 from trends in missing data. A separate lexical analysis was conducted on the
55 responses from each of the four text message inquiries. Lexical analysis is an
56 algorithm designed to automatically categorize written responses. The lexical
57
58
59
60

analyses were conducted using SPSS Text Analytics (IBM, 2011). This program used linguistics-based text analysis, which combines phrases into a common category if the differences between phrases are the use of synonyms (e.g. practicing problems and doing problems). Some of the resulting categories were then manually combined such as practice tests and old tests. Lexical analysis also facilitates an investigation into patterns of overlap among categories that provides insight into the extent that study habits are diversified at each time point.

The end result created 18 categories from the responses across the set of four inquiries. Note these categories were created independently of the codes described in Table 3. A sizable advantage of the lexical analysis technique is the ability to demonstrate the categories and interrelations between each category in a web diagram. Web diagrams were created for each exam (Figures 3 through 6) focusing only on categories with at least five responses. The web diagram represents each category with a node, and the size of the node is proportional to the frequency of the category. The frequency of each category is indicated in parenthesis inside each node. Nodes are connected with lines that indicate the extent the connected categories were mentioned together in a response. The type of line indicates the extent the categories are shared as a proportion of the smaller node. A solid line indicates that 60% or more of the responses that were categorized by the smaller node were also present in the category in the larger node. A long dash line indicates 40% to 59% agreement, a square dotted line indicates 20% to 39% agreement and no line indicates below 20% agreement. Reviewing the web diagrams can provide insight into changes in study patterns that occurred throughout the term. For context in interpreting trends in the web diagrams, the relevant topics from each exam are shown in Table 6.

Table 6: Content on Exams

Exam	Content
Exam 1	Properties of light, electron configurations, periodic trends
Exam 2	Lewis structures, molecular shapes, polarity
Exam 3	Gas laws, thermodynamics
Final Exam	Intermolecular forces, colligative properties, cumulative exam included prior content

In the Exam 1 (Figure 3) diagram, notes, previous tests, textbook and homework are the most prominent, with PowerPoint (PPT) slides also mentioned. The links show moderate overlap among these five categories, though notably no significant overlap was found between previous tests and PowerPoint or previous tests and textbook. In the Exam 2 (Figure 4) diagram the study pattern is more concentrated on notes, previous tests and textbook with moderate overlap among almost all of the categories. In Exam 3 (Figure 5) the textbook is reduced in prominence and the studying was more focused on previous tests; also the relations among nodes are generally weaker than in Exam 2 indicating less reliance on multiple study approaches. In preparing for the cumulative Final Exam (Figure 6)

1
2
3 the use of the textbook has returned to prominence along with notes and previous
4 tests, similar to Exam 2. This diagram is also the most interconnected web
5 suggesting a stronger reliance on multiple studying techniques, possibly owing to
6 the cumulative nature of this exam.
7

8 Looking for changes across study patterns, one clear trend is the diminished
9 role of studying homework in preparation for the exams. In Exam 1, homework was
10 among the most prominent nodes, whereas in each subsequent exam it is a minor
11 node. This may describe students' perceiving a lack of relevance of the homework
12 assignments in exam preparation after the first exam. For context, the online
13 homework was always due one to two days before each of these four text message
14 inquiries so that it was likely students were working on the assignments in the time
15 frame indicated. Students could also review the homework assignments after the
16 due date. Incidentally, after the semester had completed, the instructors at the
17 setting discussed deliberately including one or two questions modified from the
18 homework assignment in each of the exams to emphasize to students the
19 importance of understanding the process of problem solving in the homework over
20 simply arriving at the correct answer. By making this change it is possible that
21 students may benefit more from engaging in the homework which would be
22 reflected in their study habits and related to their academic performance.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

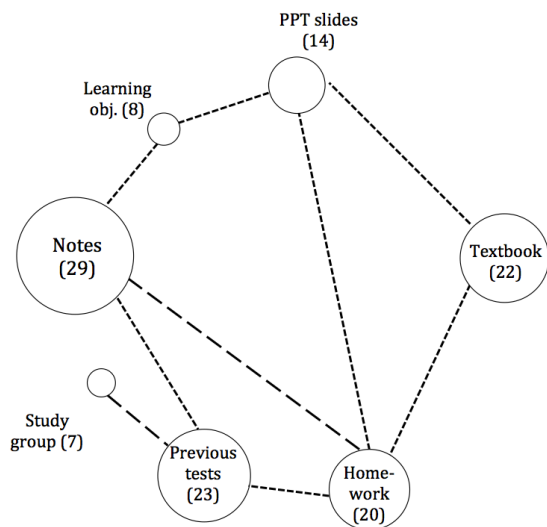


Figure 3: Exam 1 Web Diagram

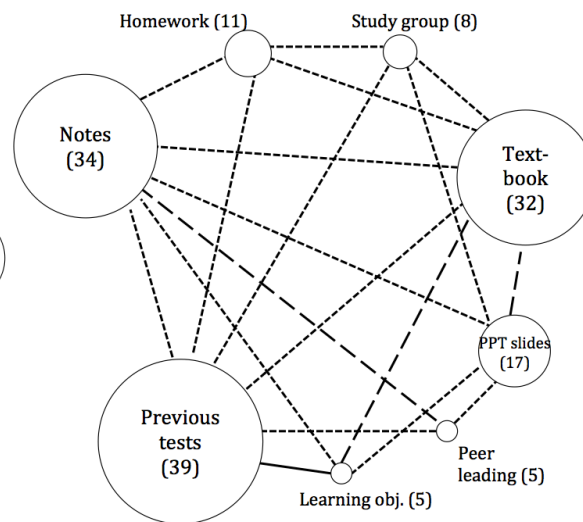


Figure 4: Exam 2 Web Diagram

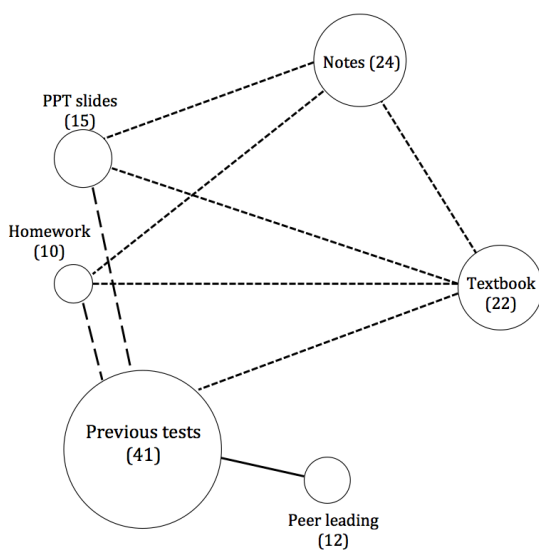


Figure 5: Exam 3 Web Diagram

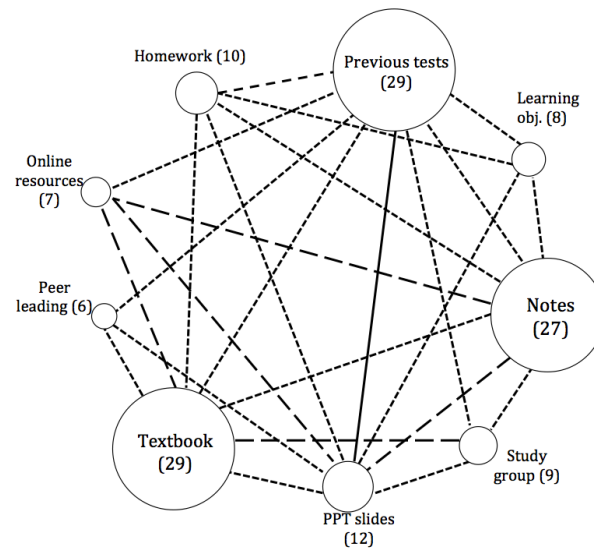


Figure 6: Final Exam Web Diagram

1
2
3
4
5 Another trend among the web diagrams is the diminished role of the
6 textbook and notes in the Exam 3 diagram. Exam 3 strongly relied on math content
7 (see Table 6) differing from the preceding exams. Students may respond to this by
8 studying the textbook and notes less and focusing more on the instructor provided
9 materials in the PowerPoint slides and previous tests. Other explanations are also
10 possible such as time constraints related to other courses giving exams, the
11 perceived challenge of the practice tests that were posted taking up more student
12 time or students finding the textbook less helpful in this content.

13
14 Returning to the research question, there appears to be considerable
15 evidence of changing study habits among a common group of students over the
16 course of the semester. The changing role of homework, textbook, notes and the use
17 of multiple study techniques suggest that student study habits differ across the
18 exams. The changes may be for many reasons including students responding to the
19 perceived effectiveness of study techniques for each exam, the perceived nature of
20 the content on each exam or the quality of study materials available, or competing
21 interests for students' time. The changes in the nature of links also indicates that
22 the variety of techniques used by students changes over the course of the semester
23 and are amplified when taking a final exam, possibly as a result of the cumulative
24 nature of the exam.
25
26
27
28

29 ***Conclusions and Future Work***

30
31 This study has shown the feasibility of using text messages to provide
32 considerable data on students' self-reported study habits in General Chemistry.
33 Among the principle limiting factors is recruiting students to participate, which may
34 become an issue depending on the intended use. Future work may benefit by
35 modifying the incentive structure for recruiting students. Second, there is evidence
36 for validity of the text messages in that the response pattern matches the expected
37 trend relative to the exam dates in the setting. Additionally, the recruited sample
38 featured minor departures from the overall population on incoming metrics,
39 including a measure of study approaches, lending support to the consideration that
40 the results are generalizable to the population of General Chemistry students at the
41 setting.
42
43

44 Next, the study provided evidence that study habits are related to academic
45 performance in the course, notably by students using study habits that are in
46 addition to the mandated course requirements. In this study, use of the textbook
47 was most prominent as the additional study habit. A direct instructional implication
48 that results from this study is the potential benefit of discussing with students the
49 need to study beyond mandatory course components. In the current study one-third
50 of the students described mandatory course components as their principal means
51 for studying and these students performed comparably to students who did not
52 report studying. Future research could have instructors discuss with students the
53 results shown here to students and measure the impact on student study habits or
54 academic performance.
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

One of the most interesting areas of future work may be an investigation on the impact of instructional techniques to impact study habits. Indeed one of the more surprising outcomes of this study was the infrequent mention of group work or studying with friends (Table 3) as it seemed possible that the weekly group work during the peer-led meetings would facilitate greater use of study groups outside of class. To investigate this area further, the impact of incorporating reform pedagogy or training sessions on how to study can be investigated in either a repeated measures or quasi-experimental design using text messages to measure students' study habits. Such a study may inform causal mechanisms behind evidence-based instructional practices. For example proponents of cooperative learning have indicated social constructivism as a potential explanation for improved learning outcomes that have been observed (Mutyhyala and Wei, 2013). The causal mechanism for learning would be that students' social processes within group-work have facilitated their conceptual understanding. An alternative causal mechanism however is that students engaged in cooperative learning may become dissatisfied with their own progress in comparison to their peers and as a result study more. The plausibility of the alternative explanation is supported by the time available outside of class relative to in class and the observed academic benefits of study habits herein. An investigation into the impact of evidence-based instructional practices on study habits can then support or dissuade the alternative explanation proposed.

This study has also shown that students' study habits can change over the course of the semester. This finding has relevance for work that relies on a single measure of particular study habits (as opposed to more general study approaches) extrapolated to describe students' habits throughout the term. It also provides some support to the expectation that instructor actions can influence student study habits. Relating changes in study habits to measures of reflective action (e.g. metacognition) or interviewing students regarding their study actions prior to each test may offer additional support for this contention. Additionally future research that investigates how changes in study habits relate to academic performance is warranted, as whether consistent or adaptable study habits are more beneficial remains an open question.

Finally, ESM has the flexibility to potentially support a diverse range of instructional strategies. For example, the action of messaging students outside of class can, by itself, serve as an instructional intervention. Instructors can use text messages to direct students toward online resources, set-up peer study groups or remind students of deadlines in a timely fashion. Additionally, the messages can be tailored for individual students or small groups; for example messages can notify a group of students who haven't completed an online homework assignment of an upcoming due date or inform a student who has struggled that the student's recent test score shows an improvement over past performance. Early research in a wide range of educational settings has shown that such tailored messages have a strong potential for producing positive gains (Dynarski, 2015). This approach may offer a non-intrusive way to show faculty concern for student performance which, when missing, has been cited as a factor in student attrition from the STEM disciplines, particularly among minority students (Tsui, 2007; Museus *et al.*, 2011).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- Aldenderfer, M. S. and Blashfield, R. K., (1984), *Cluster Analysis*, Beverly Hills, Sage Publications.
- Bernard, H. R., Killworth, P., Kronenfeld, D. and Sailer, L., (1984), The Problem of Informant Accuracy: The Validity of Retrospective Data, *Ann. Rev. Anthro.*, **13**, 495-517.
- Biggs, J. B. (2001), The revised two-factor Study Process Questionnaire: R-SPQ-2F, *Brit. J. Educ. Psychol.*, **71**, 133-149.
- Cohen, J., (1988), *Statistical Power Analysis for the Behavioral Sciences*, Hillsdale, Lawrence Erlbaum Associates.
- Cook, E., Kennedy, E. and McGuire, S. Y., (2013), Effect of Teaching Metacognitive Learning Strategies on Performance in General Chemistry Courses, *J. Chem. Educ.*, **90**, 961-967.
- Crede, M. and Kuncel, N. R., (2008), Study Habits, Skills, and Attitudes; The Third Pillar Supporting Collegiate Academic Performance, *Perspect. Psychol. Sci.*, **3**, 425-453.
- Csikszentmihalyi, M., (2014) *Applications of flow in human development and education*. Dordrecht: Springer.
- Dynarski, S. (2015), *Helping the Poor in Education: The Power of a Simple Nudge*, The New York Times, The Upshot, retrieved May 21, 2015, from: <http://www.nytimes.com/2015/01/18/upshot/helping-the-poor-in-higher-education-the-power-of-a-simple-nudge.html?smprod=nytcore-iphone&smid=nytcore-iphone-share&abt=0002&abg=1>.
- Everitt, B.S., Landau, S., Leese, M., Stah, D. (2011) *Cluster Analysis* 5th Edition. West Sussex, UK: John Wiley & Sons.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H. and Wenderoth, M. P., (2014), Active learning increases student performance in science, engineering and mathematics, *P. Natl. Acad. Sci. USA*, **111**, 8410-8415.
- Hektner, J. M., Schmidt, J. A. and Csikszentmihalyi, M., (2007), *Experience Sampling Method; Measuring the Quality of Everyday Life*, Thousand Oaks, California, Sage Publications, Inc.
- Hofmann, W., Wisneski, D.C., Brandt, M.J., and Skitka, L.J., (2014). *Morality in everyday life*. *Science*. 345, p. 1340-1343.
- IBM (2011), IBM SPSS Text Analytics for Surveys 4.0.1 User's Guide, retrieved May 21, 2015, from: ftp://public.dhe.ibm.com/software/analytics/spss/documentation/tafs/4.0.1/en/Users_Guide.pdf
- Lewis, S. E. (2011), Retention and Reform: An Evaluation of Peer-Led Team Learning, *J. Chem. Educ.*, **88**, 703-707.
- Lewis, S. E. and Lewis, J. E., (2005), The same or not the same: Equivalence as an issue in educational research, *J. Chem. Educ.*, **82**, 1408-1412.

- 1
2
3
4 Li, W.-T., Liang, J.-C. and Tsai, C.-C., (2013), Relational analysis of college chemistry-
5 major students' conceptions of and approaches to learning chemistry, *Chem.*
6 *Educ. Res. Pract.*, **14**, 555-565.
- 7
8 Museus, S. D., Palmer, R. T., Davis, R. J. and Maramba, D. C., (2011), *Factors That*
9 *Influence Success Among Racial and Ethnic Minority College Students in the*
10 *STEM Circuit*, Jossey-Bass, San Francisco.
- 11 Muthyala, R. and Wei, W., (2013), Does Space Matter? Impact of Classroom Space on
12 Student Learning in an Organic-First Curriculum, *J. Chem. Educ.*, **90**, 45-50.
- 13 Richards-Babb, M. and Jackson, J. K., (2011), Gendered responses to online
14 homework use in general chemistry, *Chem. Educ. Res. Pract.*, **12**, 409-419.
- 15 Smyth, J. M. and Stone, A.A. (2003), Ecological momentary assessment research in
16 behavioral medicine, *J. Happiness Stud.*, **4**, 35-52.
- 17
18 Tro, N. J., (2014), *Chemistry: A Molecular Approach*, Upper Saddle River, NJ, Pearson
19 Education, Inc.
- 20 Tsui, L., (2007), Effective Strategies to Increase Diversity in STEM Fields: A Review
21 of the Research Literature, *J. Negro. Educ.*, **76**, 555-581.
- 22
23 Ye, L. and Lewis, S. E., (2014). Looking for Links: Examining Student Responses in
24 Creative Exercises for Evidence of Linking Chemistry Concepts, *Chem. Educ.*
25 *Res. Pract.*, **15**, 576-586.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60