Frequent deadlines: Evaluating the effect of learner control on healthcare executives’ performance in online learning

Lawrence V. Fulton a, Lana V. Ivanitskaya b, Nathaniel D. Bastian c,*, Dmitry A. Erofeev d, Francis A. Mendez b

a Texas State University, Department of Computer Information Systems & Quantitative Methods, San Marcos, TX 78666, USA
b Central Michigan University, Department of Health Administration, Mount Pleasant, MI 48859, USA
c University of Maryland University College, Department of Mathematics, Adelphi, MD 20783, USA
d Central Michigan University, Department of Business Administration, Mount Pleasant, MI 48859, USA

Abstract

In a three-group, gender-matched, preexisting knowledge-controlled, randomized experiment, we evaluated the effect of learner control over study pace on healthcare executives’ performance in an online statistics course. Overall, frequent deadlines enhanced distribution of practice and improved learning. Students with less control over pace (in groups with weekly deadlines) spaced their study episodes to a greater extent than their peers with more control over pace (in groups with monthly and end-of-course deadlines). Online learning experience and technology self-efficacy did not explain practice distribution effects. Student perceptions of control over how, when and in which order they learn did not differ significantly across experimental groups. However, perceived control and spaced practice were positively and significantly related to performance on tests of short delayed retention and near transfer. In addition, perceived control and spaced practice predicted performance on a test of delayed retention and far transfer. Locus of control did not explain differences in performance.

1. Introduction

1.1. Background

Many course instructors can attest to their students’ dislike of frequent, instructor-set deadlines. In designing their courses, instructors must define deadline frequency, spacing and penalties for not meeting the deadlines. Once made, these decisions are often modified mid-course under pressure from learners who insist on relaxing deadlines and softening penalties. In this experimental study, we examine how middle-aged adult healthcare managers and executives distribute (or space) their practice of statistics when they are provided with different work completion deadlines. We use random assignment of learners to weekly, monthly and end-of-course deadlines. The students are enrolled in a computer-assisted (CA) distance learning course to develop their data analysis competencies. Because learner control over pace and sequence is one of the most salient features of a CA distance learning environment, there is a great need to understand how instructor-set deadlines impact distance learners.

In educational contexts, planning fallacy refers to learners’ propensity to underestimate how much time it would take to complete a learning task (Sanna & Schwarz, 2004; Schwarz, Sanna, Skurnik, & Yoon, 2007). Learners may not pay enough attention to potential obstacles and disregard scenarios that deviate from the best case. They may also overestimate how much control they have personally over when their work is completed (Newby-Clark, Ross, Buehler, Koehler, & Griffin, 2000). Multiple and frequent deadlines may improve learner awareness of planning fallacies and facilitate corrective self-regulatory actions.

Defined as “an active, constructive process in which learners plan, monitor, and control their own learning process” (Kostons, van Gog, & Paas, 2012, p. 121), self-regulated learning can be beneficial for motivating learners but it does not guarantee better outcomes, especially for novice learners (e.g., Moos & Azevedo, 2008a, 2008b; Scheiter & Gerjets, 2007; Schnackenberg & Sullivan, 2000; Steinberg, 1989). Older individuals and full-time managers are likely to be more skilled in managing deadlines than college-age students due to their greater life experience; however, their learning patterns may still be affected by course
deadlines. In a study by Ariely and Wertenbroch (2002), adult learners enrolled in an executive education program had self-regulation problems while setting and meeting deadlines. Specifically, the working professionals set overly aggressive deadlines and did not space them enough for optimal performance. As frequent and evenly spaced cues to action, instructor-set deadlines may assist learners in overcoming their self-regulation failures.

Self-determination theory (SDT) draws distinction between autonomous regulation that is integrated with one’s self (e.g., action stems from one’s interests or perceived importance) and controlled regulation in response to demands, pressures, coercions or seductions. It is unclear if instructor-set deadlines, especially frequent deadlines with penalties for late submissions, support learner autonomy (Deci & Ryan, 2006). On the one hand, deadlines can be viewed as controlling (rather than autonomy supporting) events because they impose rigorous standards for when, how often and how much the learners must study. Amabile, Dejong, and Lepper (1976) support this position by demonstrating decreases in intrinsic motivation for an interesting activity with an imposed deadline. On the other hand, the theory predicts wide variations in learners’ phenomenological experiences with deadlines. Learner A may feel pressured or coerced (“My instructor will fail me if I don’t meet a deadline”). Other learners may internalize external standards through introjections (Learner B: “If I feel bad if I don’t meet a deadline”) or identification (Learner C: “I want deadlines because I work best under pressure”). Learner D may view deadlines as an opportunity (“Learning is fun; I can explore some topics deeper if I study ahead of the recommended deadlines”). These divergent appraisals of deadlines reflect different degrees of autonomy (Ryan & Connell, 1989) and range from mostly controlled regulation by learner A to mostly autonomous regulation by learner D. According to SDT, autonomous regulation can energize learning, create a state of subjective vitality and lead to persistence. In contrast, controlled regulation leads to an inner conflict which depletes inner resources and low persistence (Deci & Ryan, 2006; Moller, Deci, & Ryan, 2006). It is unclear if instructor-set deadlines, especially frequent deadlines with penalties for late submissions, support learner autonomy (Deci & Ryan, 2006). On the one hand, deadlines can be viewed as controlling (rather than autonomy supporting) events because they impose rigorous standards for when, how often and how much the learners must study. Amabile, Dejong, and Lepper (1976) support this position by demonstrating decreases in intrinsic motivation for an interesting activity with an imposed deadline. On the other hand, the theory predicts wide variations in learners’ phenomenological experiences with deadlines. Learner A may feel pressured or coerced (“My instructor will fail me if I don’t meet a deadline”). Other learners may internalize external standards through introjections (Learner B: “If I feel bad if I don’t meet a deadline”) or identification (Learner C: “I want deadlines because I work best under pressure”). Learner D may view deadlines as an opportunity (“Learning is fun; I can explore some topics deeper if I study ahead of the recommended deadlines”). These divergent appraisals of deadlines reflect different degrees of autonomy (Ryan & Connell, 1989) and range from mostly controlled regulation by learner A to mostly autonomous regulation by learner D. According to SDT, autonomous regulation can energize learning, create a state of subjective vitality and lead to persistence. In contrast, controlled regulation leads to an inner conflict which depletes inner resources and energy (Deci & Ryan, 2006). In sum, SDT predicts that frequent deadlines may impact learners in different ways—some may be energized and others may give up—a fact that motivated the design of the present study.

The purpose of this study was to test the effects of differential learner control over pace on learners’ perceptions of control, distribution of practice, time on task and performance. Deadline frequency represents an objective measure of external control over learners’ circulatory system, blood and heart. Their analysis showed that learner control over pace but it is unclear how it translates into learners’ perceptions of control, which depletes inner resources and low persistence (Deci & Ryan, 2006; Moller, Deci, & Ryan, 2006). In contrast, learners with internal locus of control (internals) are expected to initiate behaviors that lead to desired outcomes (Deci & Ryan, 1987). Noe (1986) proposed that internals expect that their learning efforts will result in both mastery and rewards, which affects their motivation. Confirmed that motivation to learn was strongly related to locus of control and internals were more motivated than externals. To control their environment, internals engage in exploratory behaviors, for example, assess their skill gaps, goals and plans, as demonstrated by Noe and Schmitt (1986). Such self-assessment may lead to actions aimed at avoiding planning fallacies. It may also facilitate integration of instructor-set deadlines into personal learning goals, for example, when deadlines are framed as opportunities to succeed by better regulating (e.g., pacing) one’s own learning over time. Specific hypotheses and their justification are presented below.

1.2. Hypothesis testing

Hypothesis 1: In groups with frequent deadlines, learners with less control over pace will space their study episodes to a greater extent than their peers with relaxed deadlines and more control over pace. Learners prefer massing to spacing (Kornell & Bjork, 2008). Researchers studied this effect for over 100 years across a wide range of tasks most of which involved psychomotor or verbal learning. Learners prefer massing to spacing (Kornell & Bjork, 2008) even after receiving feedback that shows how their performance is hindered by massed practice. Analogous to the distribution of practice effect, a spacing effect is well documented by advertising research.
researchers interested in the temporal distribution of advertising stimuli (Janiszewski, Noel, & Sawyer, 2003). To date, five meta-analyses summarize existing research (Cepeda et al., 2006; Donovan & Radesovich, 1999; Janiszewski et al., 2003; Lee & Genovese, 1988; Moss, 1996). The distribution of practice effect can be observed in the vast majority of studies they review, e.g., 80% of studies summarized by Moss (1996). The magnitude of the distribution of practice effect varies widely as a function of multiple factors, such as study rigor (Donovan & Radesovich, 1999), the presence and the degree of spacing (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Donovan & Radesovich, 1999; Janiszewski et al., 2003; Moss, 1996), retention interval (Cepeda et al., 2006) and learning task characteristics—mental requirements and overall complexity (Donovan & Radesovich, 1999), semantic complexity (Janiszewski et al., 2003), meaningfulness (Janiszewski et al., 2003) and intentional vs. unintentional learning (Janiszewski et al., 2003). The reviewers note that less frequently studied tasks are intellectual tasks, such as computational or analytical skills used in this study, or other tasks with high mental or cognitive requirements, such as an air traffic controller simulation. Thus, despite the abundance of research, disagreement exists regarding the direction and strength of the distribution of practice effect for intellectual and complex cognitive tasks. Additional research is needed to determine if the distribution of practice effect is less pronounced for these tasks (Donovan & Radesovich, 1999; Goldstein, 1993) or if forgetting can make this effect become negative, when highly distributed practice results in poor performance (DeCecco, 1968).

1.3. Significance

Our study fills several research voids. Most of the past research on distribution of practice examined learning of youth and young adults and not middle-aged adults (Cepeda et al., 2006). Donovan and Radesovich (1999) stated that the distribution of practice effect should be tested under different boundary conditions and made a special emphasis on measuring initial level of trainee knowledge, which is what we do by including a pretest. Further, this study fills a gap in research associated with educationally-relevant Inter-Study Interval (ISI) and Retention Interval (RI), a gap that was identified in a recent meta-analysis (Cepeda et al., 2006). An ISI is the time that passes between two study episodes, typically of exactly the same content. RI is time between the last study episode and the test of what was learned. If the time is short, the test is referred to as acquisition performance. And if the time is long it is referred to as retention performance. In many studies, retention interval is only a few minutes long, but in this study it is multiple weeks long and varies by assignment. Moreover, we respond to learning researchers’ call for more research on far transfer tasks (e.g., Corbalan, Kester, & van Merriënboer, 2009) by using a delayed test of retention and far transfer—a complex data analysis assignment that integrates multiple weeks of learning material. The assignment involves cognitive tasks such as hypothesizing, testing relationships in a large dataset and drawing conclusions. It is similar to applied research problems that the learners will do in the future and is quite different from test questions used to measure near transfer. This study is likely to be of interest to educators, especially those who teach in distance learning programs, and individuals involved in the design and implementation of organizational training programs.

2. Method

2.1. Experimental design overview

The design of this research is a classical three-group, matched, pre–post, experimental design with the control group being a weekly deadlines group and the treatment groups with monthly deadlines and end-of-course deadlines. Weekly deadlines were chosen because they are common in higher education courses. An end-of-course deadline gave learners maximum possible control over pace, whereas monthly deadlines represented mid-range control. The independent, manipulated variable was learner control, an ordinal level measure based on weekly, monthly, or end-of-course deadlines. It was defined as assignment to a weekly deadlines group, monthly deadlines group and an end-of-course deadlines group. These groups mapped to low, moderate, high learner control respectively. The key difference between the design of a typical distributed practice experiment and our design is the nature of practice. Most laboratory experiments utilize simple and small tasks that can be practiced repeatedly. They may also involve continuous practice of one complex task over an extended period of time, such as an air traffic controller simulation, which requires learners to build multiple skills through trial and error. However, few laboratory experiments chunk content into multiple tasks that differ from each other and overlap only partially, as it is done in many educational settings. Educational tasks involve breaking down of complex content into small modules, practice of individual modules and then a review of a large chunk (comprised of multiple modules) to facilitate long term retention and integration of knowledge. In this study, we observed how the learners distributed their practice of multiple modules. This study received an Institutional Review Board approval.

2.2. Participants

Students enrolled in the executive education program at a comprehensive Midwestern university participated in the study. The majority of students entered the program with an expectation of improving their evidence-based, i.e., theory grounded and data driven, management practice. To that end, the students acquire theoretical foundations of sciences that support the field of health administration. They also learn hands-on research skills by conducting their own scientific investigations aimed at solving administrative problems in organizational settings. The students represented a cross-section of senior-level health administrators; about one-third of them were executives and all completed Master’s degrees prior to entering the doctoral program. The sample initially included 64 students assigned to three class sections of a 12-week statistics course designed to build learners’ skills in analyzing numerical data. Of these 64 students, only 53 were included in the study due to attrition. These working adult students had an option of enrolling in one or two courses each semester; some began with two classes and then dropped a course. Of 53 students, 55% were female. One half of the students took statistics during their first semester in the program, about 40% took it during their second year of study, and the remaining 10% took it during their third or fourth year. Estimated based on the year of awarded Bachelor’s degrees, mean age was 45 years (SD = 9.3, median = 43 years). All but five students had some experience with CA learning or computerized testing.

2.3. Random assignment

As part of the experimental design, we randomized assignment to the three groups while matching on gender. Students were neither aware that they received different deadlines until near the end of the course nor were they aware that any random assignment was conducted. At the beginning of week 11 of this 12-week course, prior to a final exam, a learner from a monthly deadline group communicated with a student from an end-of-course deadline group. As a result, he requested a relaxed deadline schedule
because monthly deadlines “put him at a disadvantage.” His request was satisfied by relaxing deadlines for all students, regardless of their experimental group. This change was unlikely to contaminate the data gathered in this study. Sufficient data was gathered after running the experiment for over 10 weeks.

2.4. Sample size, power and precision

Due to the small initial sample size of \( n = 64 \), we understood that power would be problematic for overly complex analysis. Our a priori power calculations indicated that we would be able to obtain 80% power with \( \alpha = .10 \) and an effect size of .40. Alpha was set given understanding that the sample size would be small and fixed (Table 1).

2.5. Procedure

The first and the second authors were co-instructors for all three class sections. The instructors gave learners specific dates by which they had to certify that the course module was learned at a minimally acceptable level, set at 80% correct. Within each deadline period, the learners set their study strategies, pace and module completion order. They had a paper textbook and a CA program to facilitate their learning. The CA learning program suggested a sequence of three tasks: Instruct, Practice and Certify. Learners could decide how to allocate their time among these tasks. The Instruct component had explanations, available in both text and audio, which closely followed textbook chapter parts but chunked information in the form of slides. This component reinforced textbook learning and focused on the acquisition of declarative knowledge and procedure demonstration. The Practice element built procedural knowledge through hands-on exercises and error detection, presenting trial problems similar to those that would appear on the examination. A CA tutor was available and offered step-by-step explanations, hints and correct solutions. Learners could skip one practice problem and move to the next one. Finally, the Certify requirement was used to check if a minimal mastery level was reached. Each certification attempt provided an overall pass/fail coupled with individual item error detection and detailed feedback for incorrect answers. A failed certification attempt re-directed the student back into Practice. As shown in Tables 2 and 3, successful certifications were recorded as either bonus points (+1 point for each of 45 chapter parts) or points earned toward the final grade (11 chapter reviews and chapter 14, 100 points each). No points were earned if an assignment was submitted late, which is a strong penalty.

Instruct, Practice and Certify data were automatically captured and included time on task, number of certifications, number of attempts, length of ISI and RI.

When a student re-connected to the Internet and manually uploaded a completed certificate, the other data were uploaded as well.

In all experimental groups, a weekly course outline served as a blueprint for optimal, instructor-recommended learning pace and sequence. The instructors encouraged the learners to follow this course outline closely to prepare for a data analysis assignment and two non-comprehensive examinations, one given mid-way into the course and another given at its end. Online surveys were administered at week 2 and week 8 to measure student characteristics as explained below.

2.6. Measures

Experience with online learning. The learners fell into two groups: a group of newly admitted students and a group of students who have already completed their first year of study prior to taking this statistics course. This was a nominal variable coded as “0” (first semester of online Doctoral coursework) and “1” (completed three semesters of study).

Prior knowledge. The learners were informed that they could get bonus points for participating in a pretest that measures how much

Table 2

<table>
<thead>
<tr>
<th>Module</th>
<th>Topic</th>
<th>Monthly deadline</th>
<th>Weekly deadline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–3</td>
<td>Introduction &amp; levels of measurement</td>
<td>End of m1</td>
<td>End of w1</td>
</tr>
<tr>
<td>4–7</td>
<td>Review modules 1–3</td>
<td>End of m1</td>
<td>End of w1</td>
</tr>
<tr>
<td>8–12</td>
<td>Review modules 4–7</td>
<td>End of m1</td>
<td>End of w1</td>
</tr>
<tr>
<td>13–14</td>
<td>Measures of location and dispersion</td>
<td>End of m1</td>
<td>End of w2</td>
</tr>
<tr>
<td>15–18</td>
<td>Correlation</td>
<td>End of m1</td>
<td>End of w2</td>
</tr>
<tr>
<td>19–22</td>
<td>Review modules 13–14</td>
<td>End of m1</td>
<td>End of w3</td>
</tr>
<tr>
<td></td>
<td>Review modules 15–18</td>
<td>End of m1</td>
<td>End of w4</td>
</tr>
<tr>
<td></td>
<td>Review modules 19–22</td>
<td>End of m2</td>
<td>End of w5</td>
</tr>
<tr>
<td></td>
<td>EXAMINATION 1, modules 1–18</td>
<td>Last day of w5</td>
<td>Last day of w5</td>
</tr>
</tbody>
</table>

Note. A certificate for a module review marked in bold was worth 100 grade points. A certificate for any other module was worth 1 bonus point. Examinations were not comprehensive.

Table 3

<table>
<thead>
<tr>
<th>Module</th>
<th>Topic</th>
<th>Monthly deadline</th>
<th>Weekly deadline</th>
</tr>
</thead>
<tbody>
<tr>
<td>23–27</td>
<td>The normal distribution &amp; z transformations</td>
<td>End of m2</td>
<td>End of w6</td>
</tr>
<tr>
<td></td>
<td>Review modules 23–24</td>
<td>End of m2</td>
<td>End of w6</td>
</tr>
<tr>
<td>28–29</td>
<td>Sampling distributions</td>
<td>End of m2</td>
<td>End of w7</td>
</tr>
<tr>
<td></td>
<td>Review modules 28–29</td>
<td>End of m2</td>
<td>End of w7</td>
</tr>
<tr>
<td>30–34</td>
<td>Estimating means</td>
<td>End of m2</td>
<td>End of w8</td>
</tr>
<tr>
<td></td>
<td>Review modules 30–34</td>
<td>End of m2</td>
<td>End of w8</td>
</tr>
<tr>
<td>35–38</td>
<td>Hypothesis testing means</td>
<td>End of m2</td>
<td>End of w9</td>
</tr>
<tr>
<td></td>
<td>Review modules 35–38</td>
<td>End of m2</td>
<td>End of w9</td>
</tr>
<tr>
<td>39–45</td>
<td>Hypothesis testing proportions</td>
<td>End of m2</td>
<td>End of w10</td>
</tr>
<tr>
<td></td>
<td>Review modules 39–45</td>
<td>Day 23 of m3</td>
<td>End of w10</td>
</tr>
<tr>
<td>46</td>
<td>ANOVA</td>
<td>Day 23 of m3</td>
<td>End of w11</td>
</tr>
<tr>
<td></td>
<td>DATA ANALYSIS ASSIGNMENT</td>
<td>End of w11</td>
<td>End of w11</td>
</tr>
<tr>
<td></td>
<td>EXAMINATION 2, modules 23–46</td>
<td>End of w12</td>
<td>End of w12</td>
</tr>
</tbody>
</table>

Note. A certificate for a module review marked in bold was worth 100 grade points. A certificate for any other module was worth 1 bonus point. Examinations were not comprehensive. The end-of-course deadline was day 23 of m3, which corresponds to the end of w12.
they already know about statistics and demonstrates how exami-
nations work in this course. Learners’ prior knowledge of statistics
was measured as a ratio variable and operationalized as a pretest
that was administered at the beginning of the course. The pretest
contained questions on levels of measurement, descriptive statis-
tics and probability concepts drawn from a very large item pool
that was later used for the midterm examination (see below, Tests
of shortly delayed retention and near transfer). They were asked to
do their best. The pretest was scored uniformly by the software.

Technology self-efficacy. Four items assessed technology self-
efficacy in a survey administered during week 2 (Cronbach’s alpha
= .77). All items were rated on a scale from “1 = strongly
disagree” to “6 = strongly agree.” A sample item is “I am confident
that I can learn using this educational delivery technology.”

Locus of control. Included in a week 2 survey, a locus of control
measure was adapted from Spector (1988). The external locus of
control subscale had six items, e.g., “Top scores on tests are usually
a matter of good fortune” (Cronbach’s alpha = .71). The internal
locus of control subscale had eight items, e.g., “A course is what I
make of it” (Cronbach’s alpha = .62). Response options were scaled
as follows: 1 = strongly disagree and 6 = strongly agree. A score
was calculated by subtracting the mean external locus of control
score from the mean internal locus of control score, an unweighted
linear combination. A positive value would indicate internal locus
of control, whereas a negative value would indicate a more external
locus. The range was [−5, 5].

Perceived control. Ten items adapted from Copcha and Sullivan
(2008) were used to measure control perceptions during week 8
(e.g., “In this class, I could decide how much time to spend prac-
ticing problems” and “In this class, I could decide in which order
to study different chapters”), rated on a scale from 1 = strongly
disagree to 6 = strongly agree. The items reflect learners’ autonomy
perceptions of how, when and in which order they complete
coursework (Cronbach’s alpha = .92). Such conceptualization
includes not only pace control but also other facets of control
important for learner self-regulation (Kostons et al., 2012).

Time on task. We used an automated statistics software program
(Hawkes Learning Systems), which tracks completion times,
grades, attempts, time on task (hours in Instruct, Practice and
Certify) and nearly every student interaction with the software.
The result was a robust set of data for use in this analysis. Time on task
refers to time spent in CA learning and does not include time spent
studying the textbook.

Practice distribution. The degree of practice distribution was
defined as the mean time in days that passed between the fixed
date (11 days prior to the course start) and the completion of an
assigned chapter part (bonus parts excluded), ratio level data.
One student in a weekly deadline group chose to begin her study 11
days before the course officially started. A shorter mean time from
the start date indicates that the students distribute their practice.
Students who postponed the requirements until the last hour
would have a higher mean time than those who completed the
assignment in the week the material was covered. This variable is
similar to the lay person’s notion of “cramming.”

Tests of shortly delayed retention and near transfer. Retention
performance was defined as the sum of the equally-weighted
midterm and final examinations. The two exams included a midterm
that covered material from weeks 1–5 and a final, which
covered material from weeks 6–11. Examination questions were
similar to practice questions, such as “A hospital director believes
that more than 25% of the test tubes contain errors. A sample of 160
tubes found 48 errors. Is there sufficient evidence at the .10 level to
substantiate the director’s claim?” Scored uniformly by the soft-
ware, the exams were measures of near transfer and shortly
delayed retention because they contained the same types of
problems as practice trials and were separated from the vast
majority of practice trials by at least one day’s time.

A test of delayed retention and far transfer. An SPSS-based, inte-
rative, data analysis assignment assessed far transfer and reten-
ion performance. Students’ work was graded by two co-instructors
who each graded half of assignments, after calibrating their
percentage-based scoring system using a rubric. Letter grades were
also used in analyses to ameliorate deviation associated with slight
differences in instructor grading methods. The data analysis
assignment was designed to resemble tasks faced by researchers
who work with large datasets (thousands of cases) and required
data conditioning, use of statistical analysis software, hypothesis
formulation and testing, and a write up of results interpretation
similar to what is done for research publications. This represented
far transfer because learners applied knowledge to tasks that were
functionally different from practice tasks. Students practiced
solving small, error-free and highly structured problems. In
contrast, a far transfer assignment was loosely structured and
required integration of knowledge from multiple modules to
formulate and test one’s own hypotheses from a large, uncondi-
tioned dataset with an extensive data dictionary that contained
errors. Some learners discovered that the dataset had few ratio or
interval measures that could be used to test group differences. They
developed new solutions that were not practiced in the course,
such as creating continuous variables by combining data from
multiple ordinal measures.

2.7. Data analysis

For Hypothesis 1, we used Kruskal–Wallis with post-hoc
Mann–Whitney comparisons. These non-parametric tests are
appropriate, as control over pace does not meet the assumptions
of normality and is resistant to univariate transformation. For
Hypotheses 2a and 2b, we used regression modeling on appropri-
ately transformed variables for the quantitative dependent variable
(percentage score on final examination) as estimated via Box-Cox
multivariate variable optimization, and we used multinomial
regression for the analysis associated with letter grades.

3. Results

Table 4 illustrates the descriptive statistics of the quantitative
variables included in this study and tests of group means, while
Table 5 provides the correlation matrix for untransformed quanti-
tative variables of interest. Surprisingly, the median pretest score
was only 11.43%. Significant group differences were observed for
distribution of practice (Hypothesis 1) and performance measures
(Hypotheses 2a and 2b). The three groups did not differ signifi-
cantly on pretest scores and technology self-efficacy, which are
potential confounding factors. Another potential confounder was
online learning experience, defined as newly admitted students vs.
students who completed at least one year of program coursework.
Their distribution was similar across experimental conditions
($\chi^2 = .49, p = .78, \eta = .02$). As shown in Table 4, group means for
learners’ perceptions of control were similar ($M = 5.12, SD = 5.39$
and $M = 5.04$ for students with weekly, monthly and end-of-course
deadlines, respectively, $F(2,53) = .98, p = .38$), even though
instructor control over learning pace differed greatly across groups.

An examination of intercorrelations reveals that time on task
(number of minutes spent in practice and certify) did not signifi-
cantly relate to any study variables. Therefore, it was excluded from
further analyses. Performance (exams and assignment) correlated
with all study variables, except time on task. The strongest corre-
lations were with practice distribution ($r = -.58, p < .001$ with
exam scores and $r = .55, p < .001$ with assignment scores), defined
as the average number of days from the start of the course when students completed module certification, an indicator of cramming (Hypotheses 2a and 2b).

Because practice distribution did not meet normality assumptions, we ran Kruskal–Wallis (K–W) with post-hoc Mann–Whitney tests, n = 53. The results of the K–W test suggested that experimental groups differed significantly ($\chi^2(2, N = 53) = 9.432, p < .01$). Mann–Whitney analysis of pairwise combinations identified differences among weekly vs. monthly deadlines groups (exact $p = .002$), weekly vs. end-of-course deadlines groups (exact $p = .020$), but not for monthly vs. end-of-course deadlines groups (exact $p = .838$). We found that those groups with less learner control engaged in higher levels of distributed practice than their peers with more control over pace (Hypothesis 1). The mean rank score of the weekly deadlines group (17,53) was about half that of monthly and end-of-course deadlines groups (31.88 and 31.15, respectively), a result of practical significance. Since the weekly deadlines group was different from two other groups and since monthly and end-of-course groups did not differ, we collapsed them in all further analyses. Practice distribution was not related to online learning experience (Mann–Whitney $U = 317.00, p = .54$) or technology self-efficacy, $p = -.002, p = .998$, which rules them out as potential confounders. Another potential confounder was a student in a weekly deadline group who began to study 11 days prior to the course start. After removing this student from the dataset, distribution of practice means remained in the same direction (weekly group with most distribution, followed by monthly and weekly groups) and overall group differences were statistically significant, as indicated by both a parametric test (ANOVA) and a non-parametric test (Kruskal–Wallis).

For Hypothesis 2 associated with performance on a shortly delayed test of retention and near transfer as a function of practice distribution, we estimated the following model (transformations based on Box-Cox analysis): ln(midterm + final) = intercept + b1 $\times$ pretest1/2 + b2 $\times$ weekly deadline group + b3 $\times$ monthly deadline group + b4 $\times$ ln(distributed practice) + b5 $\times$ locus of control + b6 $\times$ perceived control + error. Stepwise regressions revealed that only perceived control and distributed learning were significantly related to performance but locus of control was not. The surviving statistically significant model ($F(2,49) = 22.94, p < .001$) captured 46.25% of the variance (adjusted $R^2$) with a small standard error (.116) in a reasonable model expressed as follows: ln(midterm + final) = 3.434 $-$ .688 $\times$ ln(distributed practice) $+$ .001 $\times$ perceived control3 (see Table 6). Since the coefficient of practice distribution is negative, as indicated by both a parametric test (ANOVA) and a non-parametric test (Kruskal–Wallis).

### Table 4
Descriptive statistics and tests of group differences.

<table>
<thead>
<tr>
<th></th>
<th>Pretest scores</th>
<th>Technology self-efficacy</th>
<th>Locus of control$^a$</th>
<th>Perceived control</th>
<th>Time on task (hrs)</th>
<th>Practice distribution (days)$^b$</th>
<th>Percent scores</th>
<th>Data analysis assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Weekly Deadlines</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Mean</td>
<td>0.14</td>
<td>5.31</td>
<td>2.63</td>
<td>5.12</td>
<td>76.44</td>
<td>35.65</td>
<td>1.71</td>
<td>0.96</td>
</tr>
<tr>
<td>SD</td>
<td>0.13</td>
<td>0.70</td>
<td>0.96</td>
<td>0.76</td>
<td>27.66</td>
<td>3.68</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>Median</td>
<td>0.13</td>
<td>5.33</td>
<td>2.75</td>
<td>5.20</td>
<td>69.75</td>
<td>35.10</td>
<td>1.72</td>
<td>1.07</td>
</tr>
<tr>
<td><strong>Monthly Deadlines</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Mean</td>
<td>0.12</td>
<td>5.13</td>
<td>2.27</td>
<td>5.39</td>
<td>72.97</td>
<td>40.33</td>
<td>1.64</td>
<td>0.76</td>
</tr>
<tr>
<td>SD</td>
<td>0.05</td>
<td>0.79</td>
<td>0.89</td>
<td>0.68</td>
<td>24.88</td>
<td>4.80</td>
<td>0.36</td>
<td>0.28</td>
</tr>
<tr>
<td>Median</td>
<td>0.13</td>
<td>5.33</td>
<td>2.36</td>
<td>5.75</td>
<td>74.05</td>
<td>40.11</td>
<td>1.71</td>
<td>0.90</td>
</tr>
<tr>
<td><strong>End-of-course Deadlines</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>19</td>
<td>12</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Mean</td>
<td>0.10</td>
<td>5.42</td>
<td>2.41</td>
<td>5.04</td>
<td>72.36</td>
<td>41.63</td>
<td>1.46</td>
<td>0.71</td>
</tr>
<tr>
<td>SD</td>
<td>0.09</td>
<td>0.61</td>
<td>0.89</td>
<td>0.85</td>
<td>25.11</td>
<td>8.75</td>
<td>0.38</td>
<td>0.31</td>
</tr>
<tr>
<td>Median</td>
<td>0.10</td>
<td>5.67</td>
<td>2.29</td>
<td>5.10</td>
<td>71.10</td>
<td>39.80</td>
<td>1.56</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>52</td>
<td>48</td>
<td>48</td>
<td>52</td>
<td>53</td>
<td>53</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>Mean</td>
<td>0.12</td>
<td>5.31</td>
<td>2.45</td>
<td>5.17</td>
<td>73.85</td>
<td>39.32</td>
<td>1.62</td>
<td>0.81</td>
</tr>
<tr>
<td>SD</td>
<td>0.09</td>
<td>0.69</td>
<td>0.91</td>
<td>0.78</td>
<td>25.45</td>
<td>6.75</td>
<td>0.36</td>
<td>0.31</td>
</tr>
<tr>
<td>Median</td>
<td>0.12</td>
<td>5.33</td>
<td>2.49</td>
<td>5.25</td>
<td>71.30</td>
<td>37.81</td>
<td>1.68</td>
<td>0.90</td>
</tr>
<tr>
<td><strong>ANOVA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>0.75</td>
<td>0.69</td>
<td>0.59</td>
<td>0.98</td>
<td>0.13</td>
<td>4.37</td>
<td>2.61</td>
<td>0.36</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>1.39</td>
<td>1.03</td>
<td>1.25</td>
<td>0.98</td>
<td>0.17</td>
<td>9.43</td>
<td>5.25</td>
<td>7.67</td>
</tr>
<tr>
<td>$p$</td>
<td>0.48</td>
<td>0.51</td>
<td>0.56</td>
<td>0.38</td>
<td>0.88</td>
<td>0.02</td>
<td>0.08</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Note:  
$^a$ Higher scores indicate internal locus of control.  
$^b$ Smaller number of days indicates greater practice distribution.

### Table 5
Study variables: a correlation matrix.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Pretest scores</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Technology self-efficacy</td>
<td>0.14</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Locus of control$^a$</td>
<td>0.22</td>
<td>0.29$^*$</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Perceived control</td>
<td>0.16</td>
<td>0.34$^*$</td>
<td>0.29$^*$</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Time on task (hrs)</td>
<td>0.14</td>
<td>0.39</td>
<td>2.63</td>
<td>5.12</td>
<td>76.44</td>
<td>35.65</td>
<td>1.71</td>
<td>0.96</td>
</tr>
<tr>
<td>6. Practice distribution (days)$^b$</td>
<td>0.21</td>
<td>0.03</td>
<td>0.68</td>
<td>2.75</td>
<td>69.75</td>
<td>35.10</td>
<td>1.72</td>
<td>1.07</td>
</tr>
<tr>
<td>7. Midterm + final exam scores</td>
<td>0.28$^*$</td>
<td>0.25$^*$</td>
<td>0.46$^{**}$</td>
<td>0.38$^{**}$</td>
<td>0.04</td>
<td>0.02</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>8. Data analysis assignment</td>
<td>0.28$^*$</td>
<td>0.29$^*$</td>
<td>0.34</td>
<td>0.26</td>
<td>0.06</td>
<td>0.42$^{***}$</td>
<td>0.55$^{***}$</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note:  
$^p < .1, ^* p < .05, ^{**} p < .01, ^{***} p < .001.$  
$^a$ Higher scores indicate internal locus of control.  
$^b$ Smaller number of days indicates greater practice distribution.
regression were distributed normally and all variance inflation factors (VIF) were less than 1.1. The assumption of a linear mean function was met by transforming the variables.

For Hypothesis 2b, we also investigated the effect that pretest, experimental group assignment, practice distribution, locus of control and perceived control might have on a data analysis assignment, a test of delayed retention and far transfer. No significant predictors emerged for the assignment percentage scores. Next, we converted percentage scores to letter grades to account for slight (but important) differences in instructor grading methods. Specifically, we grouped Ds and Fs together (as only two Ds were in the sample), Cs and Bs together (as only one C was in the sample) and left A’s separate. Using multinomial regressions, we found that pretest, distributed practice, and learner perception of control were statistically significant predictors ($p < .09$, $p < .06$, and $p < .01$, respectively), while group assignment and external locus of control were not. The overall model was statistically significant ($\chi^2 = 23.06$, $p < .001$) and captured 45.2% of the adjusted variance (pseudo $R^2$).

High pretest scores and perception of control increased the likelihood of obtaining grades higher than D or F, while the effect of cramming exhibited the opposite association (little distribution of practice was associated with low grades). The small sample size affected the ability to differentiate between the groups more thoroughly. External locus of control was not statistically significant in the model but demonstrated appropriate directionality.

### 4. Discussion and conclusions

Adults face many demands on their time. It is particularly true of those who are full-time workers and full-time students simultaneously. These individuals must prioritize and manage numerous work and family deadlines, in addition to academic deadlines that are arguably the chief concern of a college-age student. Our findings indicate that for our population of interest (the senior, executive-level healthcare professional enrolled in a statistics course), the provision of frequent, evenly spaced deadlines results in greater practice distribution (Hypothesis 1), which consequently predicts performance on tests of retention and transfer, both near and far (Hypotheses 2a and 2b). Interestingly, students' perceptions of control (an indicator of perceived autonomy) assessed later in the course did not vary significantly by experimental condition. Group means were high, indicating that learners in all three experimental groups experienced high levels of autonomy. This is an important finding because it is easy to assume that students with frequent, externally imposed deadlines would perceive much less control than those with relaxed deadlines, which, according to SDT, results in lower motivation, performance and persistence. Our data suggest that this is a questionable assumption. Learners who had to adhere to twelve certification deadlines per semester experienced high levels of autonomy that were not significantly different from the levels of autonomy reported by their peers with one certification deadline per semester. But we caution against concluding that deadlines do not inhibit autonomy because other variables, such as instructors’ autonomy supportive practices, could have played a large role in students’ perceptions of autonomy. Frequent deadlines are also likely to prompt some learners to express their dissatisfaction with the course, instructor, or both and expose instructors to social pressures aimed at relaxing the deadlines. In our study, after discovering that peers had less frequent deadlines, a learner asked instructors to relax their requirements. This learner was upset about working late into the night to meet a deadline while another classmate had additional days to finish the same assignment. Student control over pacing increases satisfaction (Kraiger & Jerden, 2007), yet, our study and other studies (e.g., Alliger, Tannenbaum, Bennett, Traver, & Sholotand, 1997) suggest that it may also decrease their learning outcomes. What the learners like may not be what they need.

In sum, our findings supported Hypotheses 1 and 2. A response to Goldstein’s (1993) and Donovan & Radosovich’s (1999) call for additional research using complex cognitive tasks, our study shows that the distribution of practice effect is no less pronounced for complex cognitive tasks as it was for psychomotor or verbal learning tasks. A consistent negative predictor of learning was poorly distributed practice, when students crammed by completing many assignments within a short time frame. We support Granger & Levine’s (2010) argument that adult learners often fail to effectively regulate their own learning and argue that instructor-set deadlines help to overcome such regulation failures. Distribution of practice effects observed in this study cannot be explained by participants’ experience in online learning courses. As compared to other courses, online courses require more self-regulation, for example, in the form of self-instruction, self-evaluation and use of strategies (Zimmerman & Campillo, 2003). It could be predicted that students with less online learning experience would have more trouble with self-regulation than experienced online learners, which would affect their distribution of practice. In this study, this relationship was tested by comparing a group of newly admitted students to a group of students who completed at least one year of online courses (about 24 credit hours). No significant differences were observed across the two groups on perceptions of control and practice distribution. The students may still benefit from self-regulatory instruction; however, they are also likely to benefit from frequent deadlines.

Similarly, distribution of practice effects cannot be explained by participants’ technology self-efficacy, a factor that could play a role in how well the students are able to self-regulate with fewer deadlines. Most of the learners in this study had prior experience with CA learning and testing. Also, these adult students had a high degree of technology self-efficacy and their self-efficacy ratings did not significantly correlate with distribution of practice. Therefore, technology experience and self-efficacy may not be confounding factors. High levels of technology self-efficacy and prior exposure to CA learning among our study participants are an important observation that is at odds with stereotypical assumptions about adult learners’ technological savvy.

This study reveals characteristics of individuals that are associated with performance. Consistent with prior research on novice learners who benefit less from self-regulation than their more knowledgeable, skilled or experienced peers (e.g., Moos & Azevedo, 2008a, 2008b; Scheiter & Gerjets, 2007), this study confirms the importance of preexisting knowledge. A proxy for the amount of preexisting knowledge of statistics, a pretest was a positive predictor of performance on a data analysis assignment, a delayed test of retention and far transfer. The students brought knowledge from a variety of academic institutions where they earned a passing grade in one or more statistics courses. However, the pretest showed that their median long-term retention was only at an 11% level, a discouraging knowledge retention estimate. It is the percent of correctly solved problems taken from the first five weeks of most introductory statistics courses, the easiest material that involved the simplest of computations. The pretest was given in an
open-book format and test takers were free to use any materials and computations tools. These findings speak for the importance of the initial skill level and the need to improve long-term retention and transfer of data analytical skills taught in higher education statistics courses.

Perceived control significantly predicted performance on midterm and final exams. This relationship was replicated for the integrative data analysis assignment grades. Our findings indicate that individual differences in control perceptions matter and explain variance in academic performance beyond what is explained by spaced practice, whereas locus of control does not. In sum, a state variable of perceived control was significantly related to both performance measures whereas a trait variable of locus of control did not have any significant relationships with performance. This finding suggests that control-performance relationship may be better understood if control is conceptualized as a state rather than a trait characteristic and measured at the end of the learning program.

The study has several limitations. First, the instructors for this study were necessarily not blinded to group assignment, as administration of the course content required their knowledge of the criteria. The students were blinded, however, and the assignments were completely online. Second, students’ informal reports suggest that they practiced most in a CA learning environment, our time on task measure (total time in instruct, practice and certify) excluded time spent reading the textbook. A print textbook was mostly used in conjunction with or as a replacement of CA instruction. Inability to account for textbook learning time may explain why time on task did not have any significant relationships with other study variables. Third, a finding of no differences between weekly and end-of-course deadline groups may stem from misalignment of deadline dates. Perfect alignment could not be achieved due to unit differences. Months come in increments of 28–31 days and weeks are seven days long. For example, modules 35–38 had to be completed by the end of day 61 in a monthly deadline group and by the end of day 63 (last day of week 9) in the weekly deadline group. Fourth, the instructors had no influence regarding the amount of time spent by students working the homework and no communication regarding completion (other than the initial syllabus and online homework deadlines) was provided. Fifth, our study focused specifically on the study of statistics. Our findings may generalize to other quantitative and analytical knowledge domains, such as mathematics and accounting, however, this should be confirmed by additional research.

4.1. Implications for practice and research

Past research devoted greater attention to how spacing affected acquisition performance and did not pay as much attention to retention performance (Donovan & Radvnevich, 1999). Although we partially addressed this issue by including a data analysis assignment, more studies are needed to understand how deadlines impact long-term retention.

The weekly deadlines had a significant, practically meaningful impact on students’ spacing of practice. If executive students who are skilled in managing competing priorities are influenced by frequent, evenly spaced course deadlines with penalties, such deadlines may also improve other learners’ patterns of practice distribution.

The question is how tight the deadlines should be and how severe penalties should be for not meeting the deadlines. Is there a point at which the frequency of externally imposed deadlines becomes autonomy suppressing? In this study, we contrasted weekly, monthly and end-of-course deadlines, while imposing a penalty of 100% point loss. To address students’ dissatisfaction with strict requirements, it would be important to identify the least stringent deadline and penalty that promotes learning and autonomous regulation. For example, do biweekly deadlines work as well as weekly deadlines?

4.2. Concluding remarks

This research highlights the importance of CA instruction as a high-fidelity laboratory for learning research. CA distance education offers a chance to randomly assign learners to experimental conditions, a robust dataset for research purposes and many opportunities for fine-tuning education to maximize student learning. At the same time, it affords the design of studies with higher external validity than laboratory experiments, which makes it easier to generalize experimental findings to other training and educational settings.

References


learning in distance learning? Towards a science of distributed learning and training (pp. 65–90).