# The Learning Loss Effect in Genetics: What Ideas Do Students Retain or Lose after Instruction?

#### Amber Todd<sup>†\*</sup> and William Romine<sup>†</sup>

<sup>1</sup>Boonshoft School of Medicine and Department of Biochemistry and Molecular Biology and <sup>1</sup>Department of Biological Sciences, Wright State University, Dayton, OH 45435

#### ABSTRACT

Modern genetics is a relatively new domain, but it is increasingly important for students to have a firm grasp on the content, because genetic technologies are becoming more commonplace. In a previous study, we used the Learning Progression-based Assessment of Modern Genetics to assess high school students' knowledge of genetics concepts after an intensive inquiry-based genetics instructional period. Given that this type of intensive inquiry-based instruction is unique, we are now investigating how students' knowledge of genetics changes after instruction (i.e., learning loss effect). Using a six-measure longitudinal design, we found that students retained significant gains in five of the 12 constructs 18 months after instruction; scores were not significantly different than the pretest in seven of the constructs. Through a nonparametric analysis, we found that students are better able to retain mechanistic explanations in genetics than memorized details. This study uses an learning progression framework that examines what happens to genetics knowledge over an extended period of time after instruction and indicates that classroom time is better spent helping students construct mechanistic explanations of genetic phenomena as opposed to memorizing terminology.

## **INTRODUCTION**

Modern genetics is quickly becoming a part of everyday living with genetically modified organisms in the news and on grocery store shelves, prenatal genetic screenings being offered as part of routine obstetric care, and stem cell therapies on the rise (Lewis and Wood-Robinson, 2000; Gericke and Smith, 2014). Historically, genetics knowledge entailed classical or transmission genetics: the ideas that genes could be passed on to offspring and traits could be inherited. However, with the current molecular technologies, Stewart et al. (2005) describe that modern genetics literacy is being able to understand and integrate classical/transmission genetics (genetic model), how chromosomes recombine during meiosis (meiotic model), and molecular concepts of genes and proteins and their role in cells (molecular model) to form a richer more complex understanding of genetics as a whole. Though relatively new compared with other domains, modern genetics is foundational to biology and science literacy, as evidenced by two of the four life science disciplinary core ideas (DCIs) in the Next Generation Science Standards (NGSS) being directly related to modern genetics concepts (National Research Council, 2013). There have been a number of studies that demonstrate that genetics is difficult to learn and teach (e.g., Fisher, 1992; Marbach-Ad and Stavy, 2000; Wynne et al., 2001; Lewis and Kattmann, 2004; Stewart et al., 2005; Gericke and Smith, 2014), and four conceptual or learning progressions (LPs) have been developed for this domain (Roseman et al., 2006; Dougherty, 2009; Duncan et al., 2009; Elmesky, 2013). LPs are models of student knowledge described in terms of learning performances that increase in sophistication as student knowledge

Jennifer Knight, Monitoring Editor Submitted Oct 27, 2016; Revised Jul 5, 2018; Accepted Jul 11, 2018 CBE Life Sci Educ December 1, 2018 17:ar55 DOI:10.1187/cbe.16-10-0310 \*Address correspondence to: Amber Todd (amber.todd@wright.edu). © 2018 A. Todd and W. Romine. CBE-Life Sciences Education © 2018 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution-Noncommercial-Share Alike 3.0 Unported Creative Commons License (http://creativecommons.org/licenses/ by-nc-sa/3.0).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology. increases (Corcoran *et al.*, 2009; Stevens *et al.*, 2010; Rogat *et al.*, 2011). LPs can be used by educators to help inform instruction and assessment in a domain like genetics.

Several studies have examined students' understandings of modern genetics. Most have focused on classroom instructional interventions (e.g., Gelbart and Yarden, 2006; Rotbain et al., 2006; Elkund et al., 2007; Tsui and Treagust, 2007; Venville and Donovan, 2007; Marbach-Ad et al., 2008; Horwitz et al., 2010), a few have focused on developing valid assessments (e.g., Zohar and Nemet, 2002; Bowling et al., 2008; Smith et al., 2008; Shi et al., 2010; Tsui and Treagust, 2010; Couch et al., 2015), and fewer have tied student learning to LPs (e.g., Duncan and Hmelo-Silver, 2009; Duncan and Tseng, 2011; Duncan et al., 2016; Shea and Duncan, 2013; Freidenreich et al., 2011; Todd and Kenyon, 2016a; Todd and Romine, 2016). We have previously developed and validated the Learning Progression-based Assessment of Modern Genetics (LPA-MG), which contains items aligned to our revisions to Duncan's grade 5-10 genetics LP (Duncan et al., 2009), and used it to assess high school introductory biology students' understandings of modern genetics after an intensive inquiry-based ~23-calendar-week genetics instructional period containing four intervention units targeted to LP constructs (Todd et al., 2017b). We saw highly significant gains across the instructional period that correlated with instructionally meaningful changes in knowledge of genetics (Todd et al., 2017b). Acknowledging that this type of intensive inquiry-based instruction is unique and that students made large significant gains over the course of instruction, we are now investigating how students' knowledge of genetics changes months after instruction (i.e., investigating the learning loss effect). Learning loss can be defined as "any specific or general loss of knowledge and skills or ... reversals in academic progress, most commonly due to extended gaps or discontinuities in a student's education" (Learning loss, 2017, p. 1). Summer learning loss is the most common example, but learning loss can also occur as a result of interrupted education, returning dropouts, senior year reduced course loads, absences, ineffective teaching, or course scheduling, among other reasons (Learning loss, 2017). There are a few studies that have examined the learning loss effect, but most are focused on reading and mathematics (see Cooper et al., 1996; Cooper, 2003; Entwisle et al., 2000; Downey et al., 2004) and psychology and medical education (see Custers, 2010; Direnga et al., 2015). We are aware of no studies that have examined what happens to knowledge of genetics relative to an LP over an extended period of time after instruction, despite the large literature base in genetics and importance of genetics literacy. Here we focus on the following research questions:

- 1. Does student understanding of the genetics LP constructs change significantly during and after instruction? Are high school students able to retain their increased understanding 18 months after instruction?
- 2. What are the qualitative interpretations of the changes in understanding during and after instruction with respect to the LP?

## **REVIEW OF LITERATURE**

#### Learning Loss Effect

Several studies have investigated learning loss effects, which is not surprising, given that learning loss is something that nearly

all teachers and parents instinctually know happens. However, the literature is quite disparate and can be difficult to find. Cooper et al. (1996) state that the first known study of summer learning loss was a 1906 study on student math computation speed (White, 1906), while Custers (2010) states that knowledge retention was first studied in laboratories starting in the 1880s (see Ebbinghaus, 1966). Regardless of when the first studies were done, there has been a slow trickle of articles published on this topic in the past century. Cooper et al. (1996) compiled a review and meta-analysis of learning loss effects in the K-12 arena, mainly focused on math and reading, discussing 39 different articles and selecting 13 for meta-analysis. Overall, they found that students range from making no progress over one summer to losing 1 month of skills relative to their current grade level, and they cautioned that their analysis may be optimistic. Cooper et al. (1996) then described that gender, ethnicity, and IQ did not appear to have a consistent effect on learning losses, but socioeconomic status (SES) accounted for differences seen in reading, with lower SES students showing losses and larger declines than higher SES students. Math scores appeared to not be impacted by SES status, with all students appearing to lose the same amount of math skills over the summer (Cooper et al., 1996; Cooper, 2003).

The literature used for the meta-analysis was predominantly research in math and reading (Cooper et al., 1996); this makes sense, as state tests are more frequently conducted in math and reading than in science. However, Cooper et al.'s (1996) meta-analysis did find that the largest learning losses occurred in math computation and spelling, and they postulated that this effect was due to these skills being more procedural or involving memorization of factual information. They contend that the other skills, such as math problem solving and reading comprehension, are more conceptually based, so these are less prone to loss over a summer. They highlight that these findings are consistent with literature in cognitive psychology that suggests facts and procedural skills are the most prone to loss (Cooper and Sweller, 1987; Geary, 1995). They then go on to discuss how "From these principles we can generate some predictions about how other subject areas, not tested in recent investigations, might be affected by summer break. For example, we could speculate that over summer students will tend to forget science facts but retain knowledge of scientific concepts" (Cooper et al., 1996, p. 261, emphasis added).

Custers (2010) compiled a review of medical education literature focused on basic science knowledge. This review of literature suggested that approximately two-thirds to threefourths of knowledge will be retained after 1 year, decreasing to slightly below half after 2 years. The author discusses the utility of "Ebbinghaus's curve of forgetting" (Ebbinghaus, 1966), which describes how students have large losses at small retention intervals (i.e., immediately after instruction ends) followed by leveling off and smaller losses at longer retention intervals (i.e., many years after instruction), and how the shape of this curve seems to be the same for meaningless [retention of nonsense syllables] and meaningful [retention of ideas or concepts] knowledge, though level of retention for meaningful knowledge is higher (Custers, 2010). Custers (2010) describes that most studies seem to be supportive or at least aligned with Ebbinghaus's curve, but that the timescale of the curves can be quite different.

There are also a number of studies published about learning loss specific to the physics domain. One criticism of the review by Custers (2010) is that it predominantly focused on rote knowledge instead of more conceptual knowledge (Direnga et al., 2015); instead, the physics learning loss studies address more conceptual knowledge of physics concepts and how it changes over time. Because the genetics domain also has a large amount conceptual knowledge, these studies in physics may help us begin to think about learning loss in genetics. Briefly, Francis et al. (1998), Bernhard (1999, 2000), Pollock (2009), and Deslauriers and Wieman (2011) all demonstrate that students have strong retention levels of knowledge from force and motion, electricity and magnetism, and quantum mechanics using validated instruments for each concept. They also all show evidence in one way or another that reformed courses and interactive engagement tools increase the overall amount of knowledge students have and retain. Interestingly, Pawl et al. (2012) found a discrepancy between performance on a validated instrument (Mechanics Baseline Test [MBT]) and performance on a written test similar to the final freshman exam. Students had no change on the MBT but showed significant losses on the written exam. The authors postulated that stability on the MBT is a result of an increase in mathematical skills and a concomitant loss of basic physics concepts (as shown by the decrease on the written exam score) and recommend researchers investigate gain/loss in terms of each item or concept rather than as a whole. The results of Direnga et al. (2015) add another layer of complexity to learning loss; they showed that their students actually gained knowledge after instruction had formally ended and postulated that this was a result of some students being teaching assistants after the course and also due to the structure of Germany's exams, which occur 2 to 3 months after courses end (i.e., students begin studying for exams after the classes are over).

In the specific domain of genetics, we were able to find two references to learning loss. Dauer and Long (2015) used interview data to study loss of conceptual understanding of genetic variation and how it links with organismal fitness with college students 2.5 years after model-based instruction, finding that a majority of the students had incomplete conceptual understandings. Duncan et al., (2016) investigated whether learning either Mendelian or molecular genetics supports the learning of the other; their research design included assessments in Mendelian and molecular genetics pre/post year 1 and pre/post year 2. They observed that the mean ability estimates at the beginning of year 2 were lower than the estimates at the end of year 1 for both topics and stated "students seemed to have 'forgotten' the genetics concepts they have learned in year one, only to 're-learn' them at a similar rate, and to a similar extent in year two, explaining the lack of substantive gains between the end of the first and second years" (Duncan et al., 2016, p. 457). This study provides evidence that the learning loss effect does occur in genetics, but further research specifically investigating this is needed.

The synthesis of learning loss literature across these separate domains indicates that students tend to forget conceptually based knowledge and mechanisms less rapidly than memorized facts or rote knowledge. Learning losses/gains should be broken up into individual items or concepts, because domains as a whole tend to assess multiple overlapping things, and some items/concepts may show losses and some may show gains. The role of instruction is also important for learning as reformed courses and interactive engagement in classes tend to increase the amount of student knowledge. The amount of time that has passed since students used/reviewed the content relative to assessment administration is also important to keep in mind, consistent with the "use it or lose it" adage. With these ideas in mind, we investigated students' knowledge of genetics in 12 different constructs before, during, and after an extensive inquiry-based instructional period targeted to the constructs assessed by a validated genetics instrument.

#### Genetics LP, Constructs, and Assessment

The domain of genetics has four conceptual or LPs (Roseman et al., 2006; Dougherty, 2009; Duncan et al., 2009; Elmesky, 2013). LPs describe student understanding in the form of empirically derived learning performances that are ordered in terms of increasing difficulty or sophistication. The upper bounds of LPs are achieved through targeted instruction and curricula but are not guaranteed, even with targeted instructional interventions (Corcoran et al., 2009; Stevens et al., 2010; Rogat et al., 2011). The most widely studied and used genetics LP is the Duncan et al. (2009) LP, which targets students in grades 5–10 but has been used with college students (Todd and Romine, 2016, 2017a,b; Todd et al., 2017a,b). The LP has gone through several revisions based on empirical testing (Shea and Duncan, 2013; Todd, 2013; Todd and Kenyon, 2016a), and the revision used for this study (Supplemental Table S1) currently has 12 different genetics constructs inside the multifaceted LP. Constructs include how genetic information is organized (A), how genes code for proteins (B), the roles of proteins in cells (C1), how proteins connect genes and traits (C2), how cells express different genes (D), how genetic information is passed on to offspring (E), how genes and traits are correlated (F), how DNA varies between and within species (G1), how changes to genetic information result in increased variation and drive evolution (G2), how the environment interacts with genetic information (H), how mutations are passed on to offspring (I), and that gene expression changes (J). Supplemental Table S1 outlines these 12 constructs and the learning performances for each level.

Within this LP framework of 12 constructs, we developed and validated an assessment of modern genetics, the Learning Progression-based Assessment of Modern Genetics or LPA-MG (Todd et al., 2017b). Version 1 of the instrument was validated with the population of high school students used in this study (Todd et al., 2017b), and Version 2 of the instrument was validated with a population of college students (Todd and Romine, 2016). Both versions had significant reliability and validity evidence based on the Rasch model and were able to provide useful quantitative statistics as well as informative qualitative information about student understandings relative to LP construct levels. In our previous study, we explored high school students' ideas of genetics before, during, and immediately after a 23-week intensive inquiry-based instructional period (Supplemental Table S2; Todd et al., 2017b). We found that students made highly significant gains in all 12 constructs and described these gains in relation to the LP levels. This paper presents continued work with these same students. Here, we use the LPA-MG to assess the same students' understandings of

genetics after the instructional period has ended, investigating learning loss within an LP framework at 5 months, 14 months, and 18 months after the genetics instruction described in Todd *et al.* (2017b) ended (see Supplemental Table S2). We explore which constructs, if any, saw student retention or loss of gains.

## **METHODS**

## **Conceptual Framework**

In framing the context of this study, LPs are hypothetical models about how students' ideas change within a domain-specific learning or knowledge context (Corcoran *et al.*, 2009). The term "progression" can be somewhat of a misnomer, as students can move backward as well as forward within an LP or even transition quickly through levels so that it appears as though the student may "skip" a level (Gotwals and Songer, 2010). This potentially fluid movement between defined progression levels provides an avenue for measuring learning loss and a key motivation for studying learning loss in our genetics context.

Beyond the idea that LPs exist comes the epistemological implications for their use in research on learning whether we wish to focus on learning gains, loss, or both. Foremost, LPs remain hypothetical until sufficient data are collected that confirm their structure (Shea and Duncan, 2013), and even after such evidence is collected, they are always open to falsification and/or revision, as with any theory (Popper, 1957). The research that went into developing the genetics LP upon which we based our revisions (Duncan et al., 2009) and the fact that our revised LP framework has been explored and shown to be useful in multiple contexts (Todd and Kenyon, 2016a,b; Todd and Romine, 2016, 2017a,b; Todd et al., 2017a,b) help establish the strength of the progression we use in this study, but we nonetheless treat our LP as a model that is tentative and open to revision in light of new data. Hence the conclusions derived from our use of the LP in this study are dependent upon the extensive validation work undertaken by previous research with this progression and may change in the future in light of new work with this LP.

## Context

Seventy-four students at a midwestern, suburban, public grade 6-12 science, technology, engineering, and mathematics school participated in this longitudinal study. White/non-Hispanic students make up 69.6% of the school population; 24.5% of students are economically disadvantaged; 5.1% are students with disabilities; gender distribution is relatively equal, as students are admitted through a lottery system. The general demographics of the school are consistent with the surrounding five counties that the school serves. The 10th-grade biology curriculum was predominantly developed by the 10th-grade biology teacher (Supplemental Table S2), with the notable exception of four intervention units created by our research team to target specific aspects of the genetics LP. The first intervention unit addressed constructs B, C1, C2, D, and H (see Todd and Kenyon, 2016b); the second intervention unit addressed constructs C1, C2, and H; the third intervention unit addressed constructs A, B C2, C2, E, F, and I; the fourth intervention unit addressed constructs A, B, C1, C2, D, E, F, and I (Supplemental Table S2). The teacher indicated that construct J was only "briefly discussed" during the course of instruction. The instructional period lasted ~23 calendar weeks (early October to mid-March) but contained ~16 weeks (one 50-minute period per day) of instruction targeted to the LP constructs after accounting for holidays, snow days, and other instruction during this time. More detailed information about descriptions of units, instructional time, purpose, and descriptions of the activities and their relation to the constructs can be found in Todd *et al.* (2017b) or in Supplemental Table S2. Students completed introductory biology during their 10th-grade year (2014–2015) and chemistry during their 11th-grade year (2015–2016) as required courses. Approximately half of the students took physics during their 12th-grade year (2016–2017), as it was not a required course; students had the ability to choose science electives such as anatomy and physiology, human body systems, principles of engineering, advanced engineering design, computer programming, and others in the 11th and 12th grades.

## Instrument and Data Collection

We administered the 36-item version 1 of the LPA-MG six times over 23 months: pretest before any genetics instruction (early October of 10th grade), midpoint after 7 weeks of instruction (mid-November of 10th grade), posttest after the entire ~23 week instructional duration (mid-March of 10th grade), delayed post 1 (August of 11th grade), delayed post 2 (May of 11th grade), and delayed post 3 (September of 12th grade). These data and the coding schema are described in previous work (Todd et al., 2017b). Given that we wished to investigate the learning loss effect, we chose to assess genetics understanding at instructionally relevant times after our first study: a pretest, a midpoint during instruction, an immediate posttest, a delayed posttest after the summer, a delayed posttest after a year of school, and a delayed posttest after the next summer. Because learning loss can be due to summers and/or course scheduling (i.e., not taking biology during a year of school), we wanted to assess the students frequently over the course of their high school careers. Version 1 of the LPA-MG has been previously validated with this population of students (Todd et al., 2017b) and was determined to be a reliable unidimensional instrument. The previous paper describes the validation of the instrument with this population of students using time points 1–3 used in this paper.

The LPA-MG was constructed using a genetics LP (Duncan et al., 2009) and its revisions (Shea and Duncan, 2013; Todd, 2013; Todd and Kenyon, 2016a; Todd and Romine, 2016; Todd et al., 2017b) as a framework; it assesses 12 different genetics concepts, each aligning with the constructs in the multifaceted LP (see Supplemental Table S1). The assessment was administered using Qualtrics survey software; item order and response order for each item were randomized by Qualtrics. Administration of the same assessment six times over the course of nearly 2 years can certainly be a cause for concern; however, each item had five to seven responses from which to choose; some responses contained similar wording, given that they were tied to the LP framework, in which levels reflected increasingly sophisticated concepts and not necessarily "right" or "wrong" answers; and the items and responses were all randomized by Qualtrics, so we had measures in place to prevent students from being able to "memorize" the most advanced answers on the assessment. Students completed each administration of the assessment in a class for participation credit. Time spent on the assessment ranged from ~10 minutes to ~1 hour and 15 minutes on each administration; most students completed the pretest in ~15 minutes, the midpoint in ~30 minutes, posttest 1 in ~30 minutes, posttest 2 in ~20 minutes, posttest 3 in ~20 minutes, and posttest 4 in ~15 minutes. Of the 74 students participating in the study, 52 completed the assessment at all six time points for construct A, and 51 completed all time points for constructs B–J. Students who did not participate in all six administrations were excluded from the analysis.

#### **Statistical Analysis**

An ordered multiple-choice (OMC) scoring structure (Briggs *et al.*, 2006) was used to generate scores for students that aligned with their progression level on each item (see Supplemental Table S1; responses were aligned to levels in each construct). Due to the fact that effects of guessing are inflated on OMC items (i.e., a guessing student could randomly guess a "5" response with equal probability of guessing a "1" response), we invoked a certainty of response index requiring students to state their confidence in their item response. Students who indicated that they guessed the answer were assigned a "0" score. Previous work with the LPA-MG and high school students indicates that this correction for guessing improves test reliability and validity (Todd *et al.*, 2017b).

Each construct was represented by three items. On construct A, the combined response on the three items was used to assign students to a progression level. However, on constructs B-J, students' progression levels were estimated as averages of their scores on the three items. Items for each construct have five to seven responses that align with the respective levels of that construct, and most items within a construct have the same number of responses (constructs B, C1, E, F, G1, G2, H, I, J). Three constructs (A, C2, D) have items with differing numbers of responses; for example, construct C2 has item V10 with seven responses, while items V11 and V12 have six each. Item V10 contains a response that maps to a level of 0 on that construct, while items V11 and V12 do not contain a response that maps to a level 0, because a level 0 response did not make sense for those specific question stems (note: students could still achieve a score of 0 for those items if they indicated they were guessing). For calculating a student's progression level for items B-J, a student scoring a 3, 4, and 3 on the three respective items for a single construct would have a construct score represented as a 3.33. While we acknowledge that we could have taken the median (3) or mode (3) of a student's responses to generate a score, doing so would disregard the ability to score a 4 on one of the items. The score of a 3.33 indicates that a student has a solid level 3 understanding and is beginning to understand level 4 ideas.

We used Friedman's test in SPSS v. 21.0 as a nonparametric test to compare distributions in each of the 12 genetics concepts assessed across the six time points. This was chosen over repeatedmeasures analysis of variance methods due to the fact that our ordinal data tended to be distributed nonnormally. Students with missing data were excluded from the analysis. In this procedure, we evaluated the null hypothesis of no change in students' mean rank scores across time within each construct. Upon rejection of these univariate null hypotheses, statistical significance of students' mean rank scores between each time point were evaluated using Wilcoxon signed-rank tests with the Bonferroni adjustment for type 1 error inflation. We report SPSS Bonferroni-adjusted significance values; all null hypotheses were evaluated at the 0.05 alpha level.

#### RESULTS

## Does Student Understanding of the Genetics LP Constructs Change Significantly during and after Instruction? Are High School Students Able to Retain Their Increased Understanding 18 Months after Instruction?

The null hypothesis of no difference in mean ranks was rejected at the 0.05 alpha level for all constructs with the exception of G2 (test statistic = 10.4, df = 5, p = 0.065). SPSS Bonferroni-corrected comparisons of medians on the pretest (early October of 10th grade, time 1) and the posttest (mid-March of 10th grade, time 3) indicated that students made significant gains in 10 of the 12 constructs over the course of the instructional period (Figure 1 and Table 1; see Table 1 legend for more details). The median scores for constructs G2 (changes to genetic information result in increased variation and can drive evolution) and I (only mutations in gametes can be passed on to offspring) were not statistically different, indicating that students did not significantly increase their understanding of these concepts during instruction. These findings are consistent with our previous paper, which validated the LPA-MG using this same time 1 and time 3 data (Todd et al., 2017b). While we found significant growth across the three time points as a whole in Todd et al. (2017b), analyzing the constructs individually demonstrated that, while the students overall made significant gains in their genetics knowledge over the instructional period (and in 10 of the 12 constructs), the median scores for constructs G2 and I were not statistically different.

Relative to the time point immediately after instruction (mid-March of 10th grade, time 3), at 18 months after instruction ended (September of 12th grade, time 6), only construct A showed a statistically significant decrease in student medians (Figure 1 and Table 1).

If "retention" is defined as no statistical difference in medians between the immediate posttest (mid-March of 10th grade, time 3) and the posttest 18 months after instruction ended (September of 12th grade, time 6), then students "retained" their same level of understanding in 11 of the 12 constructs (Figure 1 and Table 1). However, this definition of retention seems to overly simplify what we see with the data. Looking at Figure 1, we see that most constructs show a parabolic curve increasing during the course of instruction (times 1-3) and decreasing after instruction (times 3-6); this is most dramatic in construct A. If we wish to define "retention" as a statistically significant increase in medians between the pretest (early October of 10th grade, time 1) and the posttest 18 months after instruction ended (September of 12th grade, time 6), then our data show that students only "retained" the increased knowledge in five (B, C2, D, E, G1) of the 12 constructs (Figure 1 and Table 1).

## What Are the Qualitative Interpretations of the Changes in Understanding during and after Instruction with Respect to the LP?

While hypothesis tests have quantitative significance, the numbers themselves do little to describe the qualitative understanding that students retained or lost relative to the genetics concepts. Similar to a box plot without whiskers, Figure 2 displays students' medians (red dot) and the interquartile range (blue box) before instruction (early October of 10th grade, time 1), immediately after instruction (mid-March of 10th grade, time 3), and 18 months after instruction ended



FIGURE 1. Box-and-whisker plot showing distributions of student scores for each construct relative to progression level over time. Red dots represent the medians. Note: LP levels are ordinal, not continuous.

(September of 12th grade, time 6) overlaid relative to the LP levels of each construct. We wish to note that LP levels are ranked and ordinal; the data representation of Figure 2 may lead readers to believe the levels are continuous, but this is not the case, which is why we took a nonparametric approach to hypothesis testing. The placement of the median and quartiles compared with the LP descriptions is meant to qualitatively show where the students fell compared with the descriptive levels so readers can better understand how student understanding of the different constructs changes across time in an instructionally meaningful way. Most constructs show moderate increases during the instructional period, moving from a less complex idea to a more mechanistically complex idea (Figure 2, time 1 to time 3), and then a small decrease 18 months after instruction (Figure 2, time 3 to time 6).

Students retained increased knowledge in constructs B, C2, D, E, and G1. Using Figure 2 and Table 2 as guides, we

TABLE 1.	Medians and	d differences	in medians	between	time	points
----------	-------------	---------------	------------	---------	------	--------

can see that construct B centers around the idea that genes code for proteins. Before instruction, the median student LP level was 2.67, corresponding to a strong understanding of the idea that genes contain information (level 2; Table 2, item V4: "DNA codes for instructions for your body") and beginning to understand that genes instruct the body at different levels (level 3; Table 2, item V4: "DNA codes for your cells, tissues, and organs"). During the course of instruction, students learned how genetic information is translated into proteins (level 6), but immediately after instruction (time 3), the median LP level was 4.67, corresponding to a strong understanding of the idea that genes code for entities inside the cell (level 4; Table 2, item V4: "DNA codes for things inside of your cells") and beginning to understand that genes code for proteins (level 5; Table 2, item V4: "DNA codes for proteins"). Thus, the median score immediately after instruction was less than the maximum for that construct, but the idea was still

				Median difference <sup>a</sup>		
Construct	Early October of 10th grade median (time 1)	Mid-March of 10th grade median (time 3)	September of 12th grade median (time 6)	Time 1 to time 3	Time 3 to time 6	Time 1 to time 6
A	2.00	5.00	2.00	-3.00***	3.00***	0.00
В	2.67	4.67	4.00	-2.00***	0.67	-1.33***
C1	2.17	3.67	4.00	-1.50***	-0.33	-1.83
C2	2.33	4.33	4.00	-2.00***	0.33	-1.67*
D	2.33	4.00	3.00	-1.67**	1.00	-0.67**
E	2.00	3.00	3.00	-1.00***	0.00	-1.00*
F	2.83	3.67	3.33	-0.84**	0.34	-0.50
G1	2.00	3.33	3.00	-1.33***	0.33	-1.00**
G2	2.33	3.33	3.00	-1.00	0.33	-0.67
Н	2.33	4.00	3.33	-1.67***	0.67	-1.00
Ι	2.33	3.00	3.00	-0.67	0.00	-0.67
J	1.83	2.67	2.67	-0.84*	0.00	-0.84

Pending rejection of the joint null hypothesis of no difference between time points (*W*, *df*, and *p* values reported in Figure 1), statistical significance of students' mean rank scores between each time point was evaluated using Wilcoxon signed-rank test with the Bonferroni adjustment for type 1 error inflation. SPSS Bonferroni-adjusted significance values are reported.

<sup>a</sup>Post hoc corrected \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001.



FIGURE 2. Condensed description of genetics LP levels with box plots showing distribution of students' scores for each construct relative to level descriptions over time. Red dots represent the medians. Blue boxes represent the first to third quartiles. Note: LP levels are ordinal, not continuous. more complex than the median idea level on the pretest. Eighteen months after instruction (time 6), the median score decreased relative to immediately after instruction, and the median LP level was 4.00, corresponding to the idea that genes code for entities inside the cell (level 4). This median score was a decrease, but it was not statistically significant. The median score 18 months after instruction (time 6) was also still significantly larger than the pretest (time 1) median score; students were able to make and retain progress on this construct, going from the introductory knowledge that genes contain information and beginning to understand that genes instruct the body at different levels, to understanding how genes code for entities inside of the cell and beginning to understand that genes code for proteins immediately after instruction, to retaining the idea that genes are instructions for the body at different levels (cells, tissues, organs, etc.) 18 months after instruction.

In construct C2, the median student moved from an understanding that changing genes changes cells before instruction (Figure 2, time 1; Table 2, item V11: "The DNA change tells the cells to change shape") to an understanding of how changes to genes change proteins to change traits immediately after instruction (time 3; Table 2, item V11: "The DNA change alters a protein which causes the change in the cell's shape") and 18 months after instruction (time 6).

In construct D, the median student moved from understanding that cells are different because they have different functions before instruction (Figure 2, time 1; Table 2, item V15: "Genes in cells are activated for the cell's function"), to an understanding that different cells have different proteins for their functions immediately after instruction (time 3; Table 2, item V15: "Only certain genes are inside each cell, producing proteins needed"), to an understanding that DNA tells the cells to be different 18 months after instruction (time 6; Table 2, item V15: "Specific genes inside the cells tell them to be different specialized cells").

In construct E, the median student moved from an understanding that offspring get half of their DNA from each parent before instruction (Figure 2, time 1) to understanding that alleles are randomly assorted immediately after instruction (time 3) and 18 months after instruction (time 6). In construct G1, the median student moved from an understanding that organisms have different DNA before instruction (Figure 2, time 1) to understanding that organisms within a species have different DNA immediately after instruction (time 3) and 18 months after instruction (time 6). Figure 2 qualitatively illustrates how student knowledge changed over ~23 months for each of the 12 constructs relative to the LP.

#### DISCUSSION

Given the context of an intensive inquiry-based ~23-calendarweek genetics instructional period containing four intervention units targeted to LP constructs assessed, it was not surprising that students made significant gains in 10 of 12 constructs over the course of instruction (time 1 to time 3). The purpose of this paper was to describe students' modern genetics knowledge relative to the LP constructs after instruction ended to understand how genetics content knowledge is affected by the learning loss effect. While only one construct showed a statistically significant decrease between the posttest immediately after

TABLE 2.	Sample	LPA-MG	version	1 items <sup>a</sup>
----------	--------	--------	---------	----------------------

Item number	Construct	Item text
V4	B: Genes code for proteins	<ul> <li>Why is DNA sometimes called the "genetic code"?</li> <li>A. DNA is your genes. (construct level 1)</li> <li>B. DNA codes for instructions for your body. (construct level 2)</li> <li>C. DNA codes for your cells, tissues, and organs. (construct level 3)</li> <li>D. DNA codes for things inside of your cells. (construct level 4)</li> <li>E. DNA codes for proteins. (construct level 5)</li> <li>F. DNA sequences code for specific amino acids. (construct level 6)</li> </ul>
V11	C2: Proteins connect genes and traits	<ul> <li>Sickle cell anemia is caused by a single letter change in the DNA being changed from an A to a T. How does this letter change cause red blood cells to change shape from round to sickle-shaped?</li> <li>A. The DNA change alters the instructions to give the sickle cell trait. (construct level 1)</li> <li>B. The DNA change tells the cells to change shape. (construct level 2)</li> <li>C. The DNA change alters a protein and the cell changes shape. (construct level 3)</li> <li>D. The DNA change alters a protein which causes the change in the cell's shape. (construct level 1) to a T.</li> <li>E. The DNA change alters an amino acid in the protein to change the cell's shape. (construct level 4) H</li> <li>F. The DNA change alters an amino acid and functionality of a protein to change the cell's shape. (construct level 6)</li> </ul>
V15	D: Cells express different genes	<ul> <li>Which of the following best describes how expression of genes is regulated or controlled?</li> <li>A. The expression of genes is not regulated or controlled. (construct level 0)</li> <li>B. Specific genes are activated in certain parts of the body. (construct level 1)</li> <li>C. Genes in cells are activated for the cell's function. (construct level 2)</li> <li>D. Specific genes inside the cells tell them to be different specialized cells. (construct level 3)</li> <li>E. Only certain genes are inside each cell, producing proteins needed. (construct level 4)</li> <li>F. All cells have the same genes but different proteins. (construct level 5)</li> <li>G. Cells activate certain genes to produce different proteins. (construct level 6)</li> </ul>

<sup>a</sup>These three items are from version 1 of the LPA-MG. Entries show item numbers, construct to which the item is aligned, text of item stem, and responses aligned with the learning progression levels. Adapted with permission from Todd *et al.*, 2017b.

instruction (time 3) and the posttest 18 months after instruction ended (time 6), pretest medians were only statistically different from the 18-month delayed-posttest medians in five constructs, indicating that students were only able to retain the gains they had made in fewer than half of the constructs. Retention in fewer than half of the constructs is in line with the review by Custers (2010), though noting the important difference that the author was discussing overall knowledge retention of a topic of less than half after 2 years rather than retention of half of the concepts assessed within a topic.

## Differences between Constructs Retained

## and Constructs Lost

Nearly all of the constructs showed a similar pattern of students moving to more mechanistically complex ideas (explaining how or why something happened; see Duncan and Tseng, 2011; Berland et al., 2016) after instruction, with a small decrease 18 months after instruction (see Figure 2). The five constructs in which students were able to retain increases were how genes code for proteins (B), how proteins connect genes and traits (C2), how cells express different genes (D), how genetic information is passed on to offspring (E), and how DNA varies between and within species (G1). Three of these constructs are more molecular in nature (B, C2, and D), one construct details meiosis (E), and one construct describes how DNA can vary and be conserved (G1). Construct B was included in five of the seven units completed during the instructional period; construct C2 in six, construct D in three, construct E in three, and construct G1 in one (Supplemental Table S2; Todd et al., 2017b), which indicates that it was not simply just increased classroom time that led to student retention.

Two constructs showed no statistically significant median increase after instruction: G2 (how changes to genetic information increase variation and drive evolution) and I (how mutations in gametes are passed on to offspring; Table 1, time 1 to time 3). Thus, students made no significant gain over the course of instruction in these two constructs to retain 18 months later; not surprisingly, the medians for these two constructs were not statistically different 18 months after instruction ended (Figure 1 and Table 1). These constructs were included in one and three of the seven units during the instructional period, respectively (Supplemental Table S2; Todd *et al.*, 2017b), so they were discussed; G2 was listed as a "key construct" for one of the 1.5-week units (Supplemental Table S2; Todd *et al.*, 2017b).

The five constructs in which students were not able to retain increases were genetic organization (A), how proteins work inside cells and how structure/function depends on the amino acid sequence (C1), how patterns are correlated between genes and traits (F), how the environment interacts with genetic information (H), and how gene expression can change at any point during an organism's life (J). These constructs were included in none (construct J) to seven (construct C1) of the seven units completed by the students (Supplemental Table S2; Todd et al., 2017b), so most were discussed over the course of the instructional period. There appear to be no overarching characteristics of these units, because they deal with concepts such as protein structure/function (C1), gene expression (J), classical genetics (F), environmental influence on genetics (H), genetic changes/variation/evolution (G2), and genetic organization (A).

Construct D displayed a significant difference between the medians on the pretest and the posttest 18 months after instruction (Table 1, time 1 to time 6), though this median difference

was quite small (-0.67)—less than one level on the progression. Constructs G2 and I displayed the same median difference between these two time points, and constructs C1, H, and J all had larger median differences, but all of these differences were statistically insignificant. With these same or larger median differences, how does one construct exhibit significant content retention (construct D) while the others exhibit no significant retention (C1, G2, H, I, J)? Because statistical significance is dependent upon the distribution of student scores as well as the median, student scores for construct C1, G2, H, I, and J were more variable than the scores for construct D (Figure 1), making it more difficult to estimate the population median precisely.

That students retained significant gains in five of the constructs 18 months after instruction ended is a promising finding. We do wish to emphasize that our instructional context was quite unique, in that the students had an intensive inquiry-based instructional period with intervention units targeted to specific LP constructs; this type of instruction is not typical in most schools. However, given that most constructs showed a similar pattern of significant increases over the course of instruction and then insignificant decreases after instruction with a mix of significant or insignificant total gains, this division of significant retention in certain constructs versus the others may be an artifact of when the assessment was given. This does not mean that the retention was not real, but rather that, if we set the cutoff at 12 or 24 months after instruction, we likely would see more or fewer constructs, respectively, have significant retention. Rather than two distinct categories of constructs that were "retained" or "lost" after instruction, we see this as a sliding continuum in which students will lose content knowledge over time (consistent with Ebbinghaus's curve), and the ones "lost" at 18 months after instruction ended were simply the first ideas to be lost. We would hypothesize that, given enough time (possibly even 6 more months), we would see all of these constructs have no significant difference between the pretest and delayed posttest; however, we acknowledge there may be certain constructs that continue to show significant gains with respect to the pretest for a substantial time post instruction.

The idea that students will forget the content they gained over the course of instruction and revert back to a base-level knowledge at a certain point after instruction is instructionally disheartening, but it highlights the need for instructors to continually reference and connect material across time between and within domains. Learning within a domain should build over time, increasing in sophistication and building on students' prior knowledge. In the NGSS, important concepts in a domain are DCIs: modern genetics concepts are mainly in LS1A (structure and function), inheritance of traits (LS3A), variation of traits (LS3B), and LS4B (natural selection; National Research Council, 2013). The same DCIs cross all grade bands, meaning that, in districts that implement the NGSS, K-12 students should continually learn about inheritance of traits, variation of traits, and the other DCIs, and new teaching should reflect increasing sophistication and, ideally, build upon ideas previously learned. This type of spiral curriculum in which DCIs are being revisited in subsequent years should help students retain knowledge and reinforce concepts. Students in this study leaving high school and entering college should encounter genetics content in their introductory college biology course. Despite their knowledge loss in genetics, they may be better poised to learn about genetics in the future, because they have had instruction in this domain and have made significant progress in most of these areas while in high school.

Because there appears to be no correlation between the length of time a construct was discussed during the course and retention of the construct, it is worth discussing the level of conceptual leaps necessary to achieve the high levels of some constructs and whether some could be "easier" to retain or lose than others. In the preceding paragraphs, we have discussed how most of the constructs exhibit slow loss over time; construct A is the exception. Construct A exhibited a very different pattern compared with the other constructs: it had the highest gain over the course of the instructional period (Table 1, time 1 to time 3, median increase of 3.00) but also was the only construct to have a significant loss after the instructional period (Table 1, time 3 to time 6, median decrease of 3.00, p < 0.001). We will now discuss this construct, why it may be so different from the others, implications from this finding related to learning loss literature, and implications for instruction.

## How Is Construct A Different from the Other Constructs?

Construct A encompasses the understanding of how genetic information is organized, or more specifically, how the concepts of "gene," "DNA," "genome," "chromosome," "nucleotide/ base," and "cell" are related to one another. The lowest level of this construct (level 0) is the idea that none of these concepts are related to one another. Level 1 comprises the understanding that some/all are related, but having incorrect correlations between the concepts (i.e., that cells make up DNA). Level 2 comprises a correct correlation between any two of the terms, level 3 comprises a correct correlation between any three of the terms, and so on, with level 6 comprising correct correlations between all six concepts. Given the importance of these terms and the concept of genetic organization, this is a fundamental construct for students to understand. Students need to understand how nucleotides make up DNA, which codes for proteins inside the cell to understand the process of translation (construct B). Conceptually, though, this construct is quite different from all the others. The other constructs are much more mechanistic (i.e., they explain how or why something happened; see Berland et al., 2016), because the levels are increasingly more sophisticated explanations for a genetics phenomenon (see Supplemental Table S1, construct B: how genes code for proteins). Construct A is more about memorizing facts about how concepts are related to one another-there really is no mechanism in this construct to describe how and why these concepts are related to one another, they just are.

The finding that this construct performed so differently from the others provides evidence that students are better able to retain the mechanistic explanations for genetics phenomena than memorized facts or details. Achieving the highest learning performance for construct A required students to make correct correlations between six different concepts, a learning performance that could be achieved by rote memorization. Achieving the highest learning performances of other constructs, in contrast, requires students to understand and explain mechanistic phenomena. To achieve the highest learning performance (level 6) of construct B (genes code for proteins), students should explain how codons in the DNA are translated to specific amino acids, which are then attached to produce a protein product. Our findings align with what Cooper *et al.* (1996) hypothesized about science in relation to learning loss—"students will tend to forget science facts but retain understanding of scientific concepts" (Cooper *et al.*, 1996, p. 261)—and findings in the physics literature (e.g., Francis *et al.*, 1998; Bernhard, 1999, 2000; Pollock, 2009; Deslauriers and Wieman, 2011; Pawl *et al.*, 2012) demonstrating that students retain conceptual physics concepts. Our findings are also consistent with cognitive psychology literature demonstrating that procedural skills and facts are the most prone to loss (Cooper and Sweller, 1987; Geary, 1995).

Instructionally, our findings and those from physics education literature (e.g., Francis et al., 1998; Bernhard, 1999, 2000; Deslauriers and Wieman, 2011) support teachers spending classroom time supporting students as they construct their own knowledge of mechanistic phenomena. Our classroom context was unique, in that the teacher spent a substantial amount of time doing inquiry-based activities that had students construct mechanistic explanations of genetics phenomena. While students still forgot their knowledge gains in more than half of the constructs assessed 18 months after instruction, we hypothesize that, if they had not undergone such intensive constructivist instruction, more constructs would have behaved similarly to construct A, for which students exhibited significant and substantial loss after instruction because they simply memorized the answers only to forget them immediately after instruction ended. That students retained gains in five of the 12 constructs 18 months after instruction ended is promising, but it highlights the need to spend classroom time helping students construct and reinforce their own knowledge, so they have a better chance of retaining that knowledge and using it as a foothold for future knowledge.

## **Implications and Future Directions**

Given that this study directly assesses learning loss in genetics tied to a genetics LP, we found that our study raises some important questions and future avenues for research. Our unique context, in which students completed inquiry-based intervention units aligned to a genetics LP and then were assessed with an instrument aligned to the same LP, is certainly not typical of genetics instruction in most classrooms. A strength of the LPA-MG is that it assesses student understanding of 12 different constructs in genetics, allowing researchers and teachers to understand which genetics topics students understand more than others. Future longitudinal research using the LPA-MG instrument with district-provided curricula in different contexts (including international) would certainly further the field in understanding what genetics concepts students tend to retain. Keeping in mind that one of the key features of LPs is that achievement of the higher levels is facilitated by targeted instruction and curricula (though not guaranteed), we would expect students without the targeted instruction to perform at slightly less advanced levels than we saw in this study; although an important thing to note is that each of the 12 constructs in the LP are important ideas in the genetics domain that should likely be taught with any genetics curriculum.

As with any study, our findings are not without limitations. The cohort of students followed in this study was the same cohort of students with whom version 1 of the LPA-MG

was validated (student reliability = 0.91, item reliability = 0.96; Todd et al., 2017b). We acknowledge that a potential conflict could exist in working with the same group of students for the validation study and then for the continued use of the instrument in this study; however, a slightly modified version 2 of the LPA-MG was validated with a different population (see Todd and Romine, 2016), and it was also found to have a high reliability (student reliability = 0.86, item reliability = 0.98) in college students, indicating the instrument is highly reliable in both populations. The way in which we scored construct A relative to the other constructs may also be a limitation to this study. The three LPA-MG items for construct A must be scored together to determine whether students are able to connect all six of the concepts correctly and thus determine at which LP level the students fall. All the other constructs had three items for which we were able to independently assign a student to a LP level. The nature of construct A necessitated this during construction of the LPA-MG instrument, and both reviewers of our validation studies and those who have used our work have not indicated issues with the scheme; however, we cannot completely rule out that this difference in scoring may have contributed to the large learning loss in construct A. Additionally, though we argue that the large difference in student performance with construct A compared with the performance with the other constructs was due to the other constructs being more mechanistic in nature, and we describe how this aligns with the literature and previous studies, this study describes one instance of one construct, so further research must be conducted to determine whether this pattern holds true across multiple contexts.

#### CONCLUSION

This study uses an LP framework to specifically examine what happens to students' understanding of genetics over an extended amount of time after instruction. We found that students lost content knowledge after instruction ended, but were able to retain gains in five of the 12 constructs assessed 18 months after instruction ended. Most importantly, we discovered that students retained understanding within constructs that required mechanistic reasoning about phenomena, but tended to lose understanding within constructs that required rote memorization. Consistent with the learning loss and cognitive psychology literature, this supports the best practice of helping students to construct their own knowledge rather than memorize facts and details.

#### ACKNOWLEDGMENTS

We acknowledge Kate Cook Whitt for help with data collection and classroom implementation.

#### REFERENCES

- Berland, K. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2016). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching*, 53(7), 1082–1112.
- Bernhard, J. (1999). How Long-Lived Is Post-course Understanding of Mechanics Concepts? Retrieved October 11, 2016, from www.researchgate .net/profile/Jonte\_Bernhard/publication/242181424\_How\_Longlived \_is\_Post-Course\_Understanding\_of\_Mechanics\_Concepts/links/ 54b8e0440cf2c27adc48fed3/How-Longlived-is-Post-Course -Understanding-of-Mechanics-Concepts.pdf

- Bernhard, J. (2000). Does active engagement curricula give long-lived conceptual understanding. *Proceedings of GIREP2000: Physics Teacher Education Beyond 2000.* Barcelona, Spain, 749–752.
- Bowling, B. V., Acra, E. E., Wang, L., Myers, M. F., Dean, G. E., Markle, G. C., ... Huether, C. A. (2008). Development and validation of a genetics literacy assessment instrument for undergraduates. *Genetics*, 178, 15–22.
- Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment*, 11(1), 33–63.
- Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Education Psychology*, 79, 347–362.
- Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation of achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*(3), 227–268.
- Cooper, H. M. (2003). *Summer learning loss: The problem and some solutions*. ERIC identifier: ED475391. Champaign, IL: ERIC Clearinghouse on Elementary and Early Childhood Education.
- Corcoran, T. B., Mosher, F. A., & Rogat, A. D. (2009). Learning progressions in science: An evidence-based approach to reform. New York: Columbia University, Teachers College, Center on Continuous Instructional Improvement.
- Couch, B. A., Wood, W. B., & Knight, J. K. (2015). The molecular biology capstone assessment: A concept assessment for upper-division molecular biology students. *CBE—Life Sciences Education*, 14(1), ar10.
- Custers, E. J. F. M. (2010). Long-term retention of basic science knowledge: A review study. Advances in Health Sciences Education, 15(1), 109–128.
- Dauer, J. T., & Long, T. M. (2015). Long-term conceptual retrieval by college biology majors following model-based instruction. *Journal of Research* in Science Teaching, 52(8), 1188–1206.
- Deslauriers, L., & Wieman, C. (2011). Learning and retention of quantum concepts with different teaching methods. *Physical Review Special Topics*— *Physics Education Research*, 7(1), 010101.
- Direnga, J., Presentati, M. B., Timmermann, M. D., Brose, A., & Kautz, C. H. (2015). Does it stick? Investigating long-term retention of conceptual knowledge in mechanics instruction. Age, 26, 1. Retrieved October 11, 2016, from https://peer.asee.org/does-it-stick-investigating-long-term -retention-of-conceptual-knowledge-in-mechanics-instruction.pdf
- Dougherty, M. J. (2009). Closing the gap: Inverting the genetics curriculum to ensure an informed public. *American Journal of Human Genetics*, *85*, 6–12.
- Downey, D., von Hippel, P., & Broh, B. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review*, 69(5), 613–635.
- Duncan, R. G., Castro-Faix, M., & Choi, J. (2016). Informing a learning progression in genetics: Which should be taught first, Mendelian inheritance or the central dogma of molecular biology? *International Journal of Science and Mathematics Education*, 14(3), 445–472.
- Duncan, R. G., & Hmelo-Silver, C. E. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching*, 46(6), 606–609.
- Duncan, R. G., Rogat, A., & Yarden, A. (2009). A learning progression for deepening students' understandings of modern genetics across the 5th– 10th grades. *Journal of Research in Science Teaching*, 46(6), 655–674.
- Duncan, R. G., & Tseng, K. A. (2011). Designing project-based instruction to foster generative and mechanistic understandings in genetics. *Science Education*, 95(1), 21–56.
- Ebbinghaus, H. (1966). Über das Gedächtnis. Untersuchungen zur experimentellen psychologie. Leipzig: Nachdruck der Ausgabe, 1885. Amsterdam, Netherlands: E. J. Bonset. Retrieved October 11, 2016, from https://psychclassics.yorku.ca/Ebbinghaus/index.htm
- Eklund, J., Rogat, A., Alozie, N., & Krajcik, J. (2007). Promoting student scientific literacy of molecular genetics and genomics. Paper presented at: Annual Meeting of the National Association for Research in Science Teaching (New Orleans, LA).
- Elmesky, R. (2013). Building capacity in understanding foundational biology concepts: A K–12 learning progression in genetics informed by research on children's thinking and learning. *Research in Science Education*, 43(3), 1155–1175.

- Entwisle, D., Alexander, K., & Olson, L. (2000). Summer learning and home environment. In Kahlenberg, R. (Ed.), *A notion at risk: Preserving public education as an engine for social mobility* (pp. 9–30). New York: Century Foundation Press.
- Fisher, K. M. (1992). Improving high school genetics instruction. In Smith, M. U., & Simmons, P. E. (Eds.), *Teaching genetics: Recommendations and research proceedings of a national conference* (pp. 24–28). Cambridge, MA: National Science Foundation.
- Francis, G. E., Adams, J. P., & Noonan, E. J. (1998). Do they stay fixed? *Physics Teacher*, *36*, 488–490.
- Freidenreich, H. B., Duncan, R. G., & Shea, N. (2011). Exploring middle school students' understanding of three conceptual models in genetics. *International Journal of Science Education*, 33(17), 2323–2349.
- Geary, D. C. (1995). Reflections of evolution and culture in children's cognition. American Psychologist, 50, 24–37.
- Gelbart, H., & Yarden, A. (2006). Learning genetics through an authentic research simulation in bioinformatics. *Journal of Biological Education*, 40(3), 107–112.
- Gericke, N. M., & Smith, M. U. (2014). Twenty-first-century genetics and genomics: Contributions of HPS-informed research and pedagogy. In Matthews, M.R. (Ed.), International handbook of research in history, philosophy and science teaching (pp. 423–467). Dordrecht, Netherlands: Springer.
- Gotwals, A. W., & Songer, N. B. (2010). Reasoning up and down a food chain: Using an assessment framework to investigate students' middle knowledge. Science Education, 94, 259–281.
- Horwitz, P., Gobert, J. D., Buckley, B. C., & O'Dwyer, L. M. (2010). Learning genetics with dragons: From computer-based manipulative to hypermodels. In Jacobson, J. J., & Reimann, P. (Eds.), *Designs for learning environments for the future: International perspectives from the learning sciences* (pp. 61–85). New York: Springer-Verlag.
- Learning loss (2017, February 23). In S. Abbott, (Ed.), *The glossary of educational reform*. Retrieved October 20, 2016, from https://edglossary.org/ learning-loss
- Lewis, J., & Kattmann, U. (2004). Traits, genes, particles and information: Re-visiting students' understandings of genetics. *International Journal of Science Education*, 26(2), 195–206.
- Lewis, J., & Wood-Robinson, C. (2000). Genes, chromosomes, cell division and inheritance—do students see any relationship? *International Journal* of Science Education, 22(2), 177–195.
- Marbach-Ad, G., Rotbain, Y., & Stavy, R. (2008). Using computer animation and illustration activities to improve high school students' achievement in molecular genetics. *Journal of Research in Science Teaching*, 45(3), 273–292.
- Marbach-Ad, G., & Stavy, R. (2000). Students' cellular and molecular explanations of genetic phenomena. *Journal of Biological Education*, 34(4), 200–205.
- National Research Council. (2013). Next generation science standards: For states, by states. Washington, DC: National Academies Press.
- Pawl, A., Barrantes, A., Pritchard, D. E., & Mitchell, R. (2012). What do seniors remember from freshman physics? *Physical Review Special Topics— Physics Education Research*, 8(2), 020118.
- Pollock, S. J. (2009). Longitudinal study of student conceptual understanding in electricity and magnetism. *Physical Review Special Topics—Physics Education Research*, 5(2), 020110.
- Popper, K. (1957). Philosophy of science. In Mace, C. A. (Ed.). British philosophy in the mid-century. London: George Allen and Unwin.
- Rogat, A., Anderson, C. A., Foster, J., Goldberg, F., Hicks, J., Kanter, D., ... Wiser, M. (2011). Developing learning progressions in support of new science standards: A RAPID workshop series. *Consortium for Policy Research in Education*. Retrieved October 11, 2016, from https://eric .ed.gov/?id=ED536834
- Roseman, J., Caldwell, A., Gogos, A., & Kurth, L. A. (2006). Mapping a coherent learning progression for the molecular basis of heredity. Paper presented at: Annual Meeting of the National Association of Research in Science Teaching (San Francisco).
- Rotbain, Y., Marbach-Ad, G., & Stavy, R. (2006). Effect of bead and illustrations models on high school students' achievement in molecular genetics. *Journal of Research in Science Teaching*, 43(5), 500–529.

- Shea, N. A., & Duncan, R. G. (2013). From theory to data: The process of refining learning progressions. *Journal of the Learning Sciences*, 22, 7–32.
- Shi, J., Wood, W. B., Martin, J. M., Guild, N. A., Vicens, Q., & Knight, J. K. (2010). A diagnostic assessment for introductory molecular and cell biology. CBE-Life Sciences Education, 9, 453–461.
- Smith, M. K., Wood, W. B., & Knight, J. K. (2008). The Genetics Concept Assessment: A new concept inventory for gauging student understanding of genetics. *CBE–Life Sciences Education*, 7, 422–430.
- Stevens, S. Y., Delgado, C., & Krajcik, J. S. (2010). Developing a hypothetical multi-dimensional learning progression for the nature of matter. *Journal* of Research in Science Teaching, 47(6), 687–715.
- Stewart, J., Cartier, J. L., & Passmore, C. M. (2005). Developing understanding through model-based inquiry (pp. 515–565). In Donovan, M. S., & Branford, J. D. (Eds.), *How students learn*. Washington, DC: National Research Council.
- Todd, A. N. (2013). The molecular genetics learning progressions: Revisions and refinements based on empirical testing in three 10th grade classrooms (Doctoral dissertation). Wright State University, Dayton, OH.
- Todd, A., & Kenyon, L. (2016a). Empirical refinements of a molecular genetics learning progression: The molecular constructs. *Journal of Research in Science Teaching*, 53(9), 1385–1418.
- Todd, A., & Kenyon, L. (2016b). How do Siamese cats get their color? *Science Teacher*, 83(1), 29–36.
- Todd, A., & Romine, W. (2016). Validation of the Learning Progression-based Assessment of Modern Genetics (LPA-MG) in a college context. *International Journal of Science Education*, *38*(10), 1673–1698.
- Todd, A., & Romine, W. (2017a). Difficulties in integrating the genetic–meiotic and genetic–molecular models of genetics. Paper presented at:

International Conference of the National Association for Research in Science Teaching (San Antonio, TX).

- Todd, A., & Romine, W. (2017b). Empirical validation of a genetics progression web for college biology students. *International Journal of Science Education*, 39(4), 488–505.
- Todd, A., Romine, W., & Correa-Mendez, J. (2017a). Modeling the transition from a phenotypic to genotypic conceptualization of genetics in a university-level introductory biology context. *Research in Science Education*. doi: 10.1007/s11165-017-9626-2
- Todd, A., Romine, W. L., & Cook Whitt, K. (2017b). Development and validation of the Learning Progression-based Assessment of Modern Genetics (LPA-MG) in a high school context. *Science Education*, 101(1), 32–65.
- Tsui, C. Y., & Treagust, D. F. (2007). Understanding genetics: Analysis of secondary students' conceptual status. *Journal of Research in Science Teaching*, 44(2), 205–235.
- Tsui, C. Y., & Treagust, D. F. (2010). Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument. *International Journal of Science Education*, 32, 1073–1098.
- Venville, G., & Donovan, J. (2007). Developing year 2 students' theory of biology with concepts of the gene and DNA. *International Journal of Science Education*, 29(9), 1111–1131.
- White, W. F. (1906). Reviews before and after vacation. *American Education*, *10*, 185–188.
- Wynne, C. F., Stewart, J., & Passmore, C. (2001). High school students' use of meiosis when solving genetics problems. *International Journal of Science Education*, 23(5), 501–515.
- Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39(1), 35–62.