**College Students' Solutions to Linear Functions Problems:
Identifying and Interpreting Common Mistakes**

BY

NOOR ELAGHA
B.A., University of Illinois at Chicago, 2017

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Master of Arts in Psychology
in the Graduate College of the
University of Illinois at Chicago, 2021

Chicago, Illinois

Defense Committee:

James Pellegrino, Chair and Advisor
Susan Goldman
Alexander Demos

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Among the domains of math knowledge that college students need to master for success in college and future careers (Conley et al., 2011), students' performance is lowest in the topic of linear functions (LF) (Britton & Henderson, 2009; Mielicki et al., 2019). Important benefits can result from analyses of performance on assessments that focus on students' ability to solve description, graph, and table LF problems and, specifically, uncovering the types of errors students make to help understand and ameliorate their struggles with this domain of mathematics. This Master's thesis project focused on exploring and modeling the cognitive underpinnings of common errors in LF problems to enable subsequent instructional research aimed at ameliorating major conceptual and/or procedural difficulties underlying performance. Empirical evidence showed that students performed considerably better on description items compared to table and graph items. Errors in the latter two problem types were mapped to the interpretation and construction aspects of the problem-solving process. Interpretation errors are predominantly observed in graph problems and errors in both interpretation and construction processes were observed for table problems. These error patterns reflected systematic inductive failures that were attributed to the overgeneralization of relevant conceptual and procedural knowledge based on familiar structural features that are conditional on the form in which the information is presented. The quantitative and qualitative analyses of the item level and student level responses can be used to inform the design of instructional interventions such as using worked examples.

I. INTRODUCTION

A. Background

Learning linear algebra in a formal instructional environment brings with it conceptual and cognitive difficulties (Dorier & Sierpinska, 2001). Understanding linear algebra requires cognitive flexibility, such as moving between tabular, algebraic, graphic, and semantic representations, as well as moving between abstract, algebraic, and geometric languages, sometimes all in one task (Alves-Dias & Artigue, 1995; Dorier 2000). Thus, it can be beneficial to evaluate the contents of student responses to problems within various linear function (LF) sub-domains, which vary based on specific features that can be found in typical algebra problems, and the conceptual and procedural knowledge and strategies needed to solve them. The present evaluation of student responses was guided by the expectations set by the Common Core State Standards Initiative (CCSI) (2010) with respect to the types of problems and specific math knowledge that college and career ready students should possess. This was done by identifying errors in the interpretation and construction aspects of the problem-solving process when solving description, graph, and table problems in order to provide insight into the cognitive underpinnings of misunderstandings that students have in this domain.

B. Linear Function Task Design

There are two principal types of actions needed to solve linear functions problems which can be classified as construction actions and interpretation actions. Construction actions comprise generation of something new, while interpretation actions comprise making sense of a functional equation, graph, or description. These actions are not mutually exclusive, and can vary in their local or global nature, as well as whether they are quantitative or qualitative (Leinhardt et al., 1990). On a cognitive level, solving LF problems involves recognizing and following relevant identification and extraction rules based on the interpretation of given information, then using the extracted parts in the process of formulating a problem model (Simon, 1978).

To address a construction action, students produce something new (Leinhardt et al., 1990; Simon, 1978). Construction actions involve building an algebraic function (including calculating and determining the elements of a function, such as slope and intercept), plotting points from a table, or building a graph. In construction actions, the generation is challenging because students must generate an answer from a part that is not given, as opposed to interpretation in which they are asked to react to given information. For example, it can be difficult to construct an equation from a given graph because it is unclear what they need to implicitly know to generate it (Markovits et al., 1986; Stein & Leinhardt, 1989; Yerushalmy, 1988). The following are Common Core standards (2010) that map to relevant construction actions in linear functions:

- Write a function that describes a relationship between two quantities. (F-BF.1)
    - Determine an explicit expression, a recursive process, or steps for calculation from a context. (F-BF.1.a)
    - Compose functions. (F-BF.1.c)
- Construct linear and exponential functions, including arithmetic and geometric sequences, given a graph, a description of a relationship, or two input-output pairs (include reading these from a table). (F-LE.2)
- Calculate and interpret the average rate of change of a function (presented symbolically or as a table) over a specified interval; estimate the rate of change from a graph. (F-IF.6)

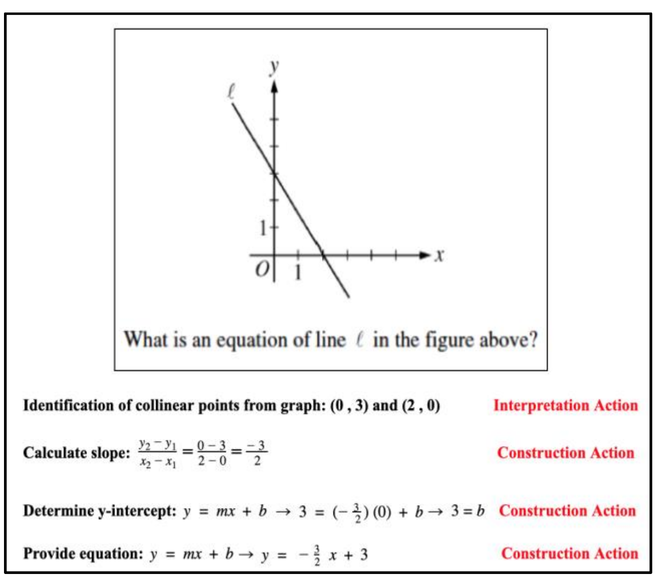
To address an interpretation action, students must first understand what the given information represents (Leinhardt et al., 1990). This information, which can be in the form of a description, graph, table, or equation, can represent a functional relationship or a situation in which a student must address a question that asks students to make meaning out of the given information. For example, interpretation actions involve interpolation (continuing a graph or table), discerning a pattern, or understanding the valence of slope in tabular and graphical form. In interpretation actions, a common struggle is a bias toward pointwise interpretations. For example, students tend to think of tables and graphs as just a compilation of individual points, instead of understanding that those points make up a conceptual entity (Schoenfeld et al., 1993; Stein et al., 1990; Yerushalmy, 1988). The following are Common Core standards (2010) that map to relevant interpretation actions in linear functions:

- Understand that a function from one set (called the domain) to another set (called the range) assigns to each element of the domain exactly one element of the range. (F-IF.1)
- For a function that models a relationship between two quantities, interpret key features of graphs and tables in terms of the quantities (F-IF.4)
- Relate the domain of a function to its graph and, where applicable, to the quantitative relationship it describes. (F-IF.5)

Linear functions questions can involve either or both interpretation and construction actions. However, it is common for construction actions to build on interpretation, but not all interpretation actions necessitate construction (Leinhardt et al., 1990). Figure 1 shows an example of a functions problem that involves both types of actions. We can learn more about students' difficulty with solving these problems by discerning the type of action for which mistakes most often occur.

Figure 1

Example Solution of Functions Problem with Stepwise Indications of the Type of Corresponding Action.



C. Significance and Nature of LF

The items studied in this project were drawn from standardized assessments and have been mapped against the Common Core State Standards for Mathematics (2010). The CCSSM specifies the levels of math knowledge at grade levels ranging from kindergarten to high school. Learning algebra and solving linear functions constitutes one of the most important proficiencies that students need for success in a variety of college majors and potential careers (Conley et al., 2011), but students struggle with these abilities for multiple reasons (Britton & Henderson, 2009; Mielicki et al., 2019). Mielicki et al. (2019) found that when testing math knowledge of college students in topics of linear equations, variables and patterns, and linear functions, students' performance was lowest for linear functions problems. Given such results, there can be a benefit to conducting a diagnostic analysis of assessments that focuses on students' ability to solve these problems and more specifically, the types of errors they make and the extent of their struggles within this mathematics domain. Two factors contribute to this struggle, task design and the preceding knowledge acquisition process; and both factors stem from the nature of the knowledge underlying the functions domain (Leinhardt et al., 1990).

Struggles with learning algebra can be attributed to the obstacle of formalism. Dorier et al. (2000) explains that because linear algebra is highly theoretical and abstract in nature, it causes what students report as feeling a "fog" (Carlson, 1997). Formalism coupled with the many new definitions and concepts that only tangentially relate to mathematical skills and concepts they already know, contribute to confusion, difficulty, and error in solving functions problems. This stems from the lack of a foundation of Cartesian geometry, which can help students to intuitively solve linear algebra problems (Robert & Robinet, 1989). Linear algebra uses abstract, algebraic, and geometric language and students must be able to understand each of these systems and move between them. Students must also intuitively understand each language, as taking these languages literally will contribute to the difficulty (Hillel, 2000). Additionally, during the knowledge acquisition process students are prone to effectively encode misinterpreted conceptual and/or procedural knowledge as a result of a discrepancy in the learning process (Ben-Zeev,1998). This often results in students applying erroneous rules to complete a problem without experiencing an impasse.

D. Challenges with LF

Piaget conducted many studies which found that the development of implicit understanding of functional logic and logical relationships begins during early childhood (Piaget et al., 1968/1977). However, learning linear algebra in a formal instructional environment brings with it conceptual and cognitive difficulties (Dorier & Sierpinska, 2001). Conceptual difficulties stem from the abstract nature of functions and linear algebra formalisms. According to Alves-Dias, understanding linear algebra requires cognitive flexibility, which entails moving between tabular, algebraic, graphic, and semantic representations, as well as moving between abstract, algebraic, and geometric languages, sometimes all in one task (Alves-Dias & Artigue, 1995; Dorier 2000).

One contribution to difficulty with functions problems is the process of operating in one representation and providing an answer in another representation, which is common in interpretation and construction actions. A reason for this is a disconnect between the understanding of equations and their application in a graphical context (Leinhardt et al., 1990; Schoenfeld et al., 1993). Examples of this difficulty include not being able to determine the y-intercept from a given graph, the valence of coordinates from a given point on a graph, or the valence of the slope of a given line on a graph. Additionally, according to Janvier (1987), the processes of moving between the different representations are not of equal difficulty. For example, the process of moving from an equation to a graph is not as difficult as moving from graphs to their equations. This is because moving from a graph involves pattern detection, which requires conceptual thinking, while moving from an equation requires a series of steps that are comparatively more straightforward, which requires procedural knowledge (Leinhardt et al., 1990). These and the aforementioned abilities are important, as explicated by the Common Core math standards (2010), and require conceptual and procedural modes of thinking.

These modes of thinking are essential to understanding functions, but can cause hardships in solving functions problems. According to Sierpinska (2000), while doing so requires conceptual thinking, students tend to overly rely on the use of procedural knowledge. Conceptual thinking refers to consciously reflecting on the different representations of information. It specifically concerns systems of concepts (as opposed to collections of independent information and ideas), as well as engaging in reasoning by making connections within a system. Procedural knowledge is guided by practical thinking which guides physical, goal-oriented actions that facilitate the actions of conceptual thinking (Sierpinska, 2000). These different methods of thinking processes are both important to understanding functions, but the tendency for students to exclusively apply procedural knowledge is what causes erroneous reasoning in functions problems (Sierpinska, 2000). Specifically, this tendency refers to difficulty in thinking analytically. That is, when they solve these problems, they do not think beyond typical examples and representations of linear algebra, and thus only apply procedural knowledge to the task at hand. This tendency could be attributed to the overwhelming amount of new concepts and abstractions during the algebra learning process (Dorier & Sierpinska, 2001), which is an exemplification of why students experience conceptual difficulties and how cognitive flexibility is pertinent.

E. Cognitive Basis of Processing Item Types and Relevant Errors

1. Description problems

Items in which students are asked to provide an equation that represents a verbal description of a linear function require the construction of a cognitive representation of the information. Kintsch and Greeno (1985) present a processing model that explicates the procedures one engages in during this process. The model comprises two main components, (1) a propositional textbase, which is a propositional structure and situation model of the information presented, and (2) a problem model, which represents the calculation strategies required to solve the problem. These components require the knowledge of propositional framing (translating sentences to propositions), properties and relations of sets, and arithmetic operations. Essentially, for every relevant textbase in the problem, a propositional frame is created which reflects their perception of the components of the textbase for which they begin to form parts of a problem model. This is a task-specific structure that represents the conceptual understanding of given information (Kintsch & Greeno, 1985). To solve the problem, conceptual and procedural knowledge is required to form this cognitive representation accurately and correctly produce an expression that represents the information to carry out necessary interpretation and construction actions. Responses to these problems that do not demonstrate the correct execution of these actions can be further analyzed to identify discrepancies in the inductive process. Errors in either or both of these procedures suggest a misunderstanding in the respective knowledge base. On the other hand, errors in description problems could suggest miscomprehension of the text as opposed to a broad conceptual and/or procedural misunderstanding of the domain (Cummings, Kintsch, Reusser & Weimer, 1988). Examining error patterns across these types of problems can indicate whether errors are caused by systematic discrepancies or other extenuating conditions.

2. <u>Graph problems</u>

Perceptual and cognitive models are constructed while solving problems that present information in the form of a graph. Specifically, interpretive processes are executed to discern meaning from the graph, and integrative processes use semantic cues to guide information extraction from the graph (Carpenter & Shah, 1998). These processes occur in tandem as student's typically engage in the following steps, according to Lohse (1991). The initial step in graph comprehension is encoding spatial features, such as general pattern recognition, then creating a structural description representing the graph in short term memory. This prompts the graph schema, which allows one to access conceptual and procedural knowledge needed to identify and extract relevant information to ultimately interpret a graph (Lohse, 1991). Once information is extracted, relevant conceptual and procedural knowledge of functions operations are used to complete the problem.

It is imperative that students have robust conceptualizations and adaptive procedural knowledge on how to solve graph problems beyond prototypical examples. In many graph problems, similar to ones used in this study, not all relevant information is explicitly presented, requiring students to apply a deeper conceptual understanding in order to extract relevant information. Especially for a domain in which the obstacle of formalism is pervasive, a robust foundational understanding of the topic, beyond rote memorization of solution procedures and general formulaic understanding, is imperative (Schoenfeld, 1985; 1988; 1992). For example, as seen in Figure 1 shown earlier, a line could be presented on a graph with no explicitly given coordinates of collinear points which are needed for a comprehensive understanding of the graph to extract other information from the representation.

3. Table problems

Problems that display discrete information in tables contain data points which require the extraction of relevant information and the construction of a representation of the relationship between two variables (Vessey, 1991). Contrary to graphs, tables do not contain substantial spatial information or require symbolic or semantic mapping of information (Lohse, 1993; Vessey, 1991). Inductive reasoning is the process of evaluating specific information to formulate a rule that represents given premises (Johnson-Laird & Byrne,1993; Klauer, 2001). This analytic process is required to extract and use information from tables in order to detect patterns and make inferences about given information (Vessey, 1991; Neubert & Binko, 1992). In the context of LF table problems, the solution process requires specific inductive reasoning processes: pattern induction and pattern extrapolation which comprise analytic and systematic evaluation of given values then inferring the pattern of relations of the values relative to each other, respectively. These are the processes required in the interpretation and construction actions that make up the solution process to LF table problems. Responses to these problems that do not demonstrate the correct execution of these actions can be further analyzed to identify specific discrepancies in the inductive process.

F. Study Goals

The goals of this study were to: (1) determine college students performance on interpretation and construction actions required to solve LF problems and whether their performance varies across items in which information is presented in different forms, (2) observe and identify specific errors that students tend to make for each type of problem so that they can be mapped to the hypothesized cognitive processes needed to solve these problems, and (3) determine the specific errors that are most prevalent that could be addressed in an instructional intervention that aims to ameliorate them.

Students were given a series of linear functions problems that required them to demonstrate the actions they take in generating a solution. Those actions presumably call upon the understanding and application of the knowledge associated with relevant Common Core math standards as described earlier. Responses were analyzed for overall accuracy, and inaccurate responses were further analyzed for mistakes in a construction action, in an interpretation action, or general misunderstanding. Identifying sources of difficulty in solving these problems can inform subsequent research focused on interventions to remedy common mistakes that students make while solving these problems using procedures such as worked examples with and without error detection and error correction.

II. METHODS

A. Materials

The items used in this study were derived from the item pool used in the study conducted by Mielicki et al. (2019). The latter materials were drawn from multiple sources, including practice tests for the ACT exam and the College Board's SAT exam, and items that were publicly provided from the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study (TIMSS), and the Partnership for Readiness for College and Careers (PARCC). The Mielicki et al. (2019) study included 136 items that were selected based on content and construct representation relative to Common Core standards for middle-school and high-school levels, 44 of which were functions questions. The items in Mielicki et al. (2019) were presented in their original five-alternative multiple-choice format using a computer-based assessment administration platform. In the present study, items were chosen from the set of LF items used by Mielicki et al. (2019) based on their mapping to Common Core math standards for algebra but were presented in an open response format in contrast to Mielicki et al. (2019).

Appendices A, B, and C display the total set of 21 items included in the item pool along with each item's *p*-value based on results from Mielicki et al. (2019) and a companion study conducted with UIC undergraduates. The p-values for each item indicates the proportion of subjects who answered the problem correctly. The *p*-values shown for each item are based on data pooled across the two separate data collections and reflect performance based on 180-200+ undergraduates responding to each item when presented in a five-alternative multiple-choice format.

The items shown vary in the information presented and the nature of the problem students are asked to solve. Information is provided in either a graphical, tabular, or descriptive format, and the items require students to engage in interpretation and construction processes that ostensibly call upon understanding of relevant Common Core math standards. The items in Appendix A present verbal descriptive information and ask students to produce an expression representing the situation verbally described. The items in Appendix B present information in a tabular format and ask students to produce an expression that represents the relationship underlying the tabled values. The items in Appendix C present information in a graphical format and ask students to produce an expression representing the graphed data or calculate and provide the slope or intercept of the expression representing the graphed data. In Appendix C there are two items without *p*-values since these tasks were not included in the Mielicki et al. (2019) study. These tasks were chosen to be similar in requirements to the other graph related tasks and were drawn from NAEP and a practice test for the College Board's SAT exam. All the tasks in

Appendices A, B, and C require interpretation and construction actions as specified by standards included in the Common Core (2010).

All tasks were presented in the open response format shown in Appendices A, B, and C allowing students to explicitly show their knowledge of how to solve these problems. Similar to what was presented earlier in Figure 1, a majority of the steps in solving these problems require interpretation and construction actions. For each item type, mistakes in responses were categorized based on the action(s) that corresponds with each step in solving the problems within each general class. This was determined by whether the mistake indicated non-understanding of one or more of the relevant math standards.

B. Design and Procedure

Data were collected from 207 total subjects using the UIC Student Subject Pool. Of the 207 total subjects tested, there was valid item performance data from 187 subjects and problem solution data in the form of written solutions to the problems for 151 of those 187 subjects. Participants were presented with a set of 21 math problems relating to linear functions using an online survey platform, Qualtrics, in an open response format. All items were divided into 3 blocks that were counterbalanced in such a way that across subjects there were 3 unique block orders. Additionally, each block included items representing each of the 3 major item types in random order. Students were asked to solve each of the problems and to explicitly show how they arrived at their answer. Answers were coded for accuracy, and when an answer was inaccurate, the written work was subsequently analyzed for mistakes in a construction action, mistakes in an interpretation action, or general non-understanding. After completing all 21 problems, participants responded to a demographic information questionnaire (Appendix D) that included questions about age, year in school, major/career plans, scores on the ACT/SAT, math and science classes taken in college thus far and the average grade earned, and education level.

C. Scoring and Analysis

Each item was scored for overall correctness using a dichotomous 0-1 scoring procedure. For those cases where errors occurred, the nature of the error was categorized using a qualitative categorization scheme to identify construction errors, interpretation errors, or both. The error classification system is described further in the results and discussion section that follows. The distribution of error types was evaluated for each item and cluster of similar items. The open response nature of these tasks can yield ambiguous responses that are difficult to classify. For cases in which errors occurred but the type of error could not be clearly identified, responses were coded as "unclear".

III. RESULTS AND DISCUSSION

In the sections that follow, various analyses of performance are presented and the main results of each analysis are briefly discussed. The presentation begins with analyses of accuracy patterns on the various item types and then turns to analysis of specific types of errors on each problem type. The error classification analyses are then followed by analyses of the consistency of error types within and across the LF problem types. The final section examines demographic variables that predict student performance on the LF items.

A. Analyses of Item Performance

All responses were scored for accuracy and triaged to make sure that the responses were sufficient for the purpose of quantitative and qualitative analyses. Scatterplots and frequency distributions of Subject performance on each of the three classes of items showed that there were some subjects whose performance on all item types was uniformly low. A subsequent examination of their actual responses indicated that these 30 subjects were not attempting the tasks in any serious manner, including the easier Description items, and thus their data were eliminated from the dataset submitted to further analysis. The data from the remaining 151 subjects for whom written solutions were obtained were then analyzed to verify that each of the three problem sets provided evidence of internal coherence and reliability while also examining relationships between subject and item performance on each of the three item subsets.

1. Overall Accuracy

Analysis of the accuracy data for all student responses showed that the easiest items were those in which information was presented as a Description and the most difficult were those in which information was presented in a Graph. Table items also showed significant error levels. Table 1 shows the *p* values for each item which indicate the proportion of subjects who answered each problem correctly with an average *p* value for each item subset.

Table 1

Performance Data for Each Item Type

Description	***p - value***	**Graph**	***p - value***	**Table**	***p - value***
D1	0.91	**G1**	0.27	**T1**	0.46
D2	0.68	**G2**	0.41	**T2**	0.60
D3	0.69	**G3**	0.47	**T3**	0.50
D4	0.86	**G4**	0.46	**T4**	0.56
D5	0.81	**G5**	0.63	**T5**	0.63
D6	0.78	**G6**	0.42	**T6**	0.58
D7	0.77	**G7**	0.40	**T7**	0.56
Average *p*	**0.78**	**Average *p***	**0.44**	**Average *p***	**0.55**

At the subject level, overall accuracy scores were obtained for each subject on each problem type. Table 2 shows the means and variances of subject performance on each of the three item types. Correlations were then computed between each of the three sets of items using overall subject performance on each item type. The correlations were modest and in expected directions given differences in the means and variances of student performance for each item type (Table 2). As shown in Table 3, the highest correlation was between performance on Table and Graph items, $r(149) = .60$, $p < .001$).

Table 2

Subject-level Performance Data for Each Item Type

Item Type	Mean	SD
Description	5.46	1.34
Graph	3.08	2.40
Table	3.85	2.66

Note. Indicated are the mean and standard deviation of items solved correctly out of 7 total items.

Table 3

Correlations Between Performance on the Three Item Types

	Description	Graph	Table
Description	1		
Graph	0.49*	1	
Table	0.44*	0.60*	1

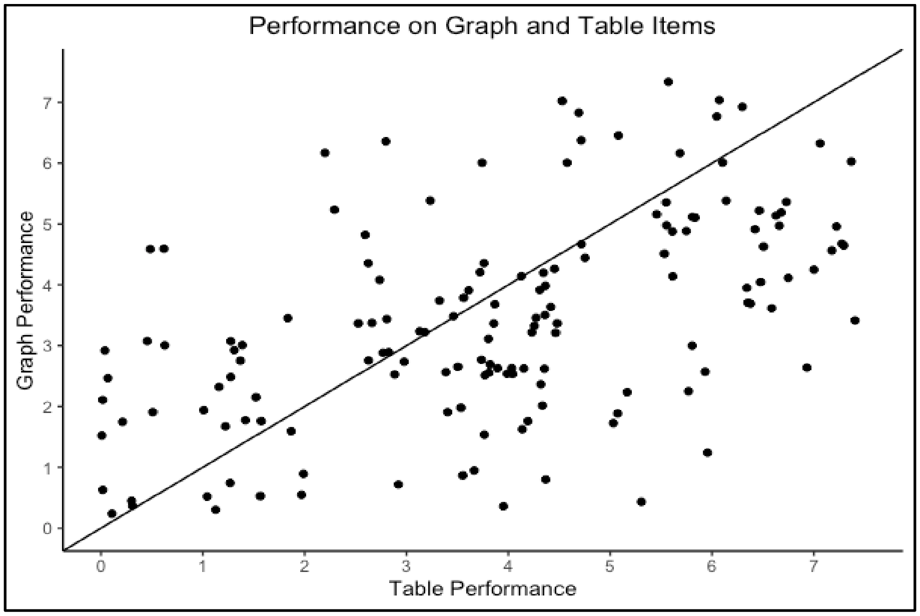
Note: *$p < .001$

Figure 2 shows a scatterplot which displays the bivariate distribution of individual students' performance across the two most difficult and highly correlated problem types as indicated above. The distribution of the points relative to the diagonal (equality of performance on each problem type) conveys how well or how poorly students performed on each of the two most difficult problem types. Individuals whose performance lies along the diagonal performed equally well or equally poorly on both item types. If their performance is below and away from the diagonal they performed better on table items than graph items. If their performance is above and away from the diagonal they performed better at graph items than table items.

A majority of the individuals fall into a cluster just below the line and around the center of the x-axis, which is consistent with results indicating that average performance on graph items (*M* = 3.08, *SD* = 2.40) was lower than performance on table items (*M* = 3.85, *SD* = 2.66). Individuals whose performance lies toward the left extremity of the diagonal generally performed somewhat better on the graph items than the table items. In contrast, individuals whose performances lies toward the right extremity of the diagonal generally performed better on the table items than the graph items. A least-squares regression line was fit to the data and is plotted in Figure 2. The regression line is represented by the equation, $y = 0.46x + 1.60$. The slope and intercept of the regression equation show a relationship between Table and Graph performance consistent with the description provided above. At the lower end of the performance scale students performed better on Graph than Table items, thus the non-zero intercept, but as performance on Table items improves performance on Graph items improves at roughly half the rate.

Figure 2

Scatterplot Depicting Distribution of Individual Students' Performance Across Graph and Table Problems



Note. Performance is represented by the number of problems they solved correctly out of 7 total problems in each item set. The regression line is represented by the equation, y = 0.46x + 1.60.

A further analysis of the highest and lowest performers was done to evaluate individual differences in performance patterns within and between these groups. To do this, the data for highest performing 27% (n = 39) and lowest performing 27% (n = 39) subjects were extracted and the accuracy patterns were compared given their performance on each item type (see Kelley, 1939). Results showed a significant difference in highest performers, $t(38) = 3.33$, $p = 0.001$, d = 1, in which they did slightly better on table problems (M = 6.03, SD = 1.05) than graph problems (M = 5.12, SD =.73). Results also showed a significant difference in lowest performers, $t(38)$ = 3.07, $p = 0.004$, d = 0.52, in which they did slightly better on graph problems (M = 2.25, SD = 1.37) than table problems (M = 1.33, SD = 2.07). Further analysis of these groups regarding specific error patterns within item type subsets is discussed in a subsequent section on *Sub-error consistency across subjects and items*.

2. Psychometric Analyses

To examine how the items within each problem type functioned relative to each other and across the entire set of items, various test theory analyses of item performance were conducted. The goal of these analyses was to establish whether item performance within problem types showed evidence of internal consistency and reliability of measurement when considered by item type and as a whole with respect to assessing knowledge of linear functions.

Table 4 shows the discriminability levels of the items within each set with respect to the other items within each given subset. As shown in Table 4 the discriminability levels were modest for the Description items and considerably higher for the Graph and Table items. Table 5 shows item discriminability levels when performance on individual items within a given problem type was compared to performance across the entire set of 21 items tested. The results again show lower values for the Description items relative to the Graph and Table items. Such results are consistent with the pattern of overall correlations of subject performance across item types.

Table 4

Discriminability of Each Item Compared to Total Performance on its Subset Membership

Description	Discrimination	Graph	Discrimination	Table	Discrimination
D1	0.25	G1	0.54	T1	0.96
D2	0.68	G2	0.64	T2	0.82
D3	0.39	G3	0.82	T3	0.93
D4	0.18	G4	0.86	T4	0.71
D5	0.32	G5	0.68	T5	0.64
D6	0.39	G6	0.79	T6	0.82
D7	0.25	G7	0.61	T7	0.68

Table 5

Discriminability of Each Item Compared to Total Performance on All Items

Description	Discrimination	Graph	Discrimination	Table	Discrimination
D1	0.25	G1	0.57	T1	0.82
D2	0.61	G2	0.61	T2	0.82
D3	0.14	G3	0.89	T3	0.82
D4	0.14	G4	0.71	T4	0.71
D5	0.29	G5	0.54	T5	0.57
D6	0.29	G6	0.71	T6	0.79
D7	0.18	G7	0.57	T7	0.64

Internal reliability analyses based on Cronbach's Alpha and alpha-if-deleted values showed modest to high internal consistency in each item subset and no particularly problematic items (Table 6). Consistent with the prior results on item discriminability, there was less internal consistency in performance among the Description items, a reasonable level of internal consistency in performance among the Graph items, and a very high level of internal consistency among the Table items.

Table 6

Alpha Values of Each Item Subset if the Item is Dropped From the Subset

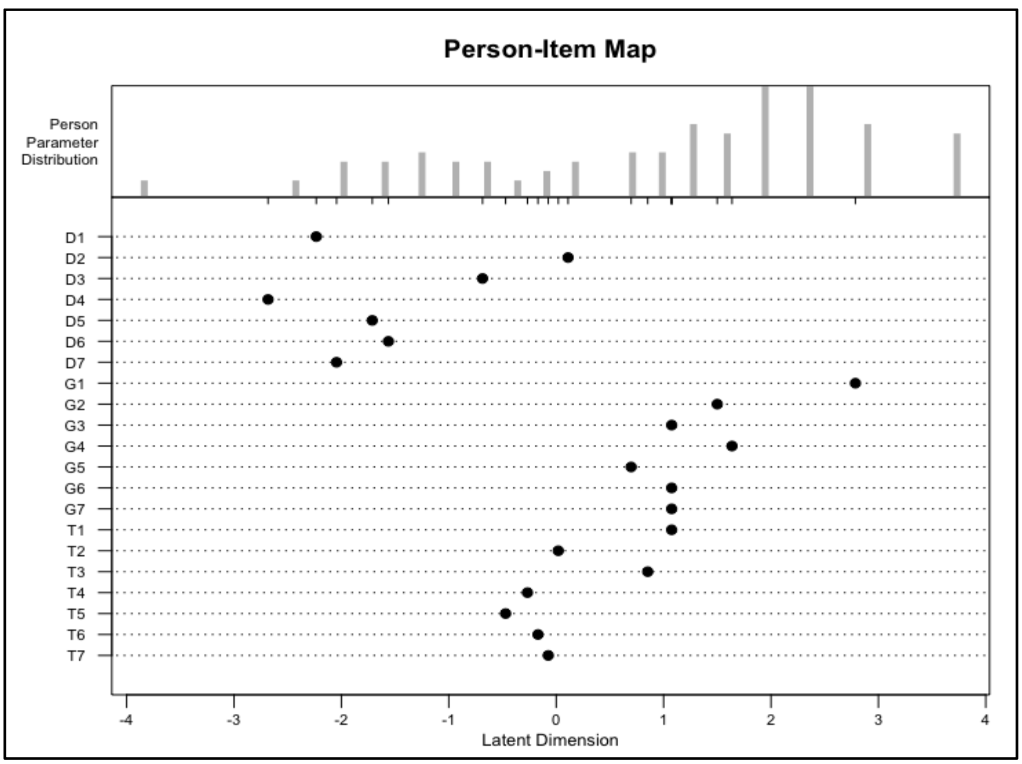
α = 0.59; 95% CI = [0.46, 0.72]		α = 0.74; 95% CI = [0.66, 0.82]		α = 0.91; 95% CI = [.89, .94]	
Description	**Alpha-if-deleted**	**Graph**	**Alpha-if-deleted**	**Table**	**Alpha-if-deleted**
D1	0.56	G1	0.73	T1	0.91
D2	0.58	G2	0.72	T2	0.90
D3	0.65*	G3	0.70	T3	0.90
D4	0.55	G4	0.70	T4	0.90
D5	0.54	G5	0.72	T5	0.91
D6	0.46	G6	0.69	T6	0.89
D7	0.54	G7	0.72	T7	0.91

Note. The asterisk (*) denotes an alpha-if-deleted value that is above the alpha value for the overall subset.

Item Response Theory analyses were also conducted by fitting a 1-parameter Item Response Theory (Rasch) model to the individual student scores on each item, with final responses to all items scored dichotomously as correct or incorrect. The Rasch analysis provides a difficulty parameter for each item as well as an ability score for each respondent represented by a relative value along the latent dimension scale. This resulted in a good distribution of item difficulties that ranged between - 2.64 and + 2.81 (Figure 3). An Andersen's Likelihood Ratio Test (Andersen, 1973), which is a goodness of fit test that detects any item bias based on the students' latent ability, was conducted. The result of Anderson LR-test ensured that the model fit the data (^{2} $\chi 2(16) = 21.35$, $p = 0.17$).

Figure 3

Rasch Person-Item Map Depicting the Distribution of Overall Performance on Each Item and Students' Ability Measure.



The results of these analyses indicate that the three classes of items are distinguishable in terms of subject performance with evidence of varying levels of coherence among the specific items within each item subset. Clearly the Description items behave differently than the Table and Graph items with respect to coherence of the items within the subset. Much of this is attributable to the significantly higher levels of accuracy on each of the Description items and the set overall. While the Description items hang together as a group and can be differentiated from performance on the Table and Graph items, errors on these items are less consistent from item to item. Even so, the set has a reasonable level of reliability as a distinct item subset should one wish to use it as a separate subtest for LF knowledge. In contrast, the Table and Graph items demonstrated substantial levels of internal coherence and both could serve as reliable subtests of specific forms of LF knowledge.

B. Analysis of Errors

A coding system was used to classify the type(s) of errors observed for each problem type. Errors can be classified as interpretation errors, construction errors, both, neither, or unclear. In the process of classifying erroneous responses, student work was analyzed to identify where and why the error occurred in the problem-solving process. When classifying an erroneous response as an interpretation error, the response must depict a lapse in the student's grasp of the conceptual underpinnings for interacting with the given information and/or a discrepancy in their cognitive representation of the given information. This is manifest in different ways based on item type because of the different conceptual and procedural knowledge required for solving each item type.

When classifying an erroneous response as a construction error, the response must depict a lapse in the conceptual and/or procedural knowledge necessary to construct a simple linear equation using the canonical y = mx + b form. Construction errors focus on a students' failure to apply relevant procedural knowledge in using the information that was extracted from the problem in order to produce a final solution expressed as a linear function. Thus, if a subject incorrectly represents one of the key elements in the problem situation, such as the slope or intercept, in their interpretation of the problem but otherwise maps those interpretations onto an equation in proper form the error is not one of construction but interpretation. This classification of what constitutes a construction error is consistent across all three item types given that they all ultimately require students to produce a simple linear equation or some transform of one. To do this they must produce a "syntactically" correct linear equation, and be able to use it to solve for the relevant unknown which requires procedural knowledge for execution.

Below are actual student responses which serve as examples of erroneous solutions to description, table, and graph problems using the general interpretation and construction coding scheme and which has been further detailed into subcategories of each error type to precisely identify the specific knowledge discrepancy in interpretation and construction actions. The subcategories will be described for each corresponding item type in the sections that follow.

C. Description Problem Errors

As shown in Table 1, description items showed higher accuracy than the graph and table items, resulting in comparatively fewer erroneous responses to classify and analyze. The majority of these were interpretation errors such that there is a misrepresentation of a relationship stated verbally in constructing the symbolic expression using the canonical y = mx + b form for a simple linear equation. Some erroneous solutions to description problems are shown below.

The Student Response to Description Item 1 (Figure 4) represents one of the few common errors in description problems. The work suggests that their propositional framing of the given information “twice as many adults" prompts multiplication knowledge leading to an interpretation of this as the variable multiplied by 2. It also suggests that their propositional framing of the given information “4 more than” also prompts multiplication knowledge leading to an interpretation of this as 2*c* multiplied by 4. It suggests that there is no propositional distinction between “twice a value” and “a value more than”, and that both textbases prompt the same knowledge of properties and relations of sets. The discrepancy in the framing yielded an erroneous interpretation of the verbal description which carried over to the constructed equation. Essentially, this response demonstrated a discrepancy in their conceptual understanding of the description of the different values and the relation of those values to the critical elements of the linear equation.

Figure 4

Example Student Response to a Description Item

Description Item 1:	**Student Response:**
(D3) A car can seat c adults. A van can seat 4 more than twice as many adults as a car can. Write an expression in terms of c that describes how many adults the van can seat.	**c = amount of adults seated in a car** **a van can sit twice the amount times 4, y = the amount of adults seated in a van** **therefore** **y=4(2c)**

Similarly, the Student Response to Description Item 2 (Figure 5) shows that their interpretation of the given information “$48 for each hour worked” was interpreted as “where the plumber starts to charge” in which they represented $48 as the additional cost; and their interpretation of the given information “plus an additional $9 for travel” as “the $9 can change depending on the location of the individual” in which they represented $9 as the value that is contingent on the number of hours worked. It suggests a discrepancy in the framing of the proposition “$48 for each hour worked” and the proposition “plus an additional $9 for travel” in their cognitive representation of the problem and the mapping then to the components of the symbolic expression. Additionally, these responses would not be considered construction errors because the students attempt to represent the relationship between the given values as a linear equation in y = mx + b form, showing that they are familiar with the essential elements of a simple linear equation and how to represent the given or extracted information.

Figure 5

Example Student Response to a Description Item

Description Item 2:	**Student Response:**
(D1) A plumber charges customers $48 for each hour worked plus an additional $9 for travel. If h represents the number of hours worked, write an expression that could be used to calculate the plumber's total charge in dollars?	**y=9h+48** **The $48 tells where the plumber starts to charge from and the $9 can change depending on the location of the individual.**

Because description items showed higher accuracy than the graph and table items, resulting in comparatively fewer erroneous responses to classify and analyze, they were excluded from further error analysis. Instead, the bulk of the error analysis effort was devoted to the Table and Graph items as described below.

D. Table Problem Coding Scheme

The following are examples of some common erroneous solutions to a table problem and their coding using the aforementioned scheme.

Student Response 1 (Figure 6) is representative of responses in which students appear to evaluate the changes in the *x* values and the *y* values separately and this is but one example of this pattern of problem solution. Student Response 2 (Figure 6) is representative of responses in which students provide a set of separate equations that only focus on describing the relationship between corresponding *x* and *y* values. These solutions suggest erroneous interpretations of the data pattern in the given table and a failure to demonstrate the necessary inductive reasoning process (pattern induction and extrapolation), which is germane to solving table problems. These responses can also be considered construction errors because they do not show evidence of attempting to represent the relationship between the *x* and *y* values as a linear equation in y = mx + b form. More generally, they show a lack of knowledge of the relationship between the pattern in the table and a mapping of the pattern to the canonical form of a linear function.

Figure 6

Example Student Responses to a Table Item

<table>
<tr>
<td rowspan="2"><u>Table Item:</u>
<table>
<tr><th>x</th><th>y</th></tr>
<tr><td>0</td><td>-3</td></tr>
<tr><td>1</td><td>-1</td></tr>
<tr><td>2</td><td>1</td></tr>
</table>
(T2) Write an equation which describes the three pairs of x and y in this table.</td>
<td><u>Student Response 1:</u>
x increases by 1
y increases by 2

x=y+2</td>
</tr>
<tr>
<td><u>Student Response 2:</u>
x-3 because 0-3=-3
X-2 because 1-2=-1
X-1 because 2-1=1</td>
</tr>
</table>

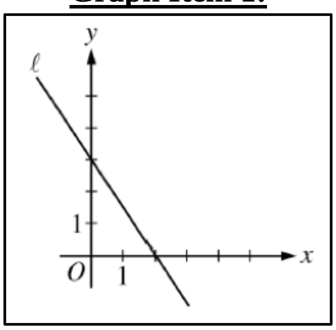
E. <u>Graph Problem Coding Scheme</u>

Some erroneous solutions to Graph problems are shown below. These solutions suggest the erroneous interpretation of the specific graph and weakness in understanding key aspects of graphs and how they map to the elements of a linear equation. Students appear to be relying on procedural knowledge or prototypical knowledge of graphs. These responses are representative of different solution variations for common errors.

The Student Response to Graph Item 1 (Figure 7) represents common demonstrations of a discrepancy in subjects' conceptual knowledge of graphs and linear functions (aside from the incomplete solution of this particular response). This response shows a discrepancy in the understanding of how the slope of a line is calculated and, more holistically, how the slope describes the properties of the line, suggesting an interpretation error. Here they used the given information to apply their broad knowledge of the slope equation - identifying the corresponding *x* and *y* intercepts as the "rise" and the "run" - to find the slope of the line, as well as not considering the valence of the line and how that would contribute to the slope. This would also be considered a type of construction error (aside from the fact that they did not complete the solution) because they identified slope as "rise over run" suggesting that they possess procedural knowledge to solve for slope, but did so incorrectly. This is a demonstration of how relying on procedural knowledge of linear functions is not sufficient for correctly solving these types of problems.

Figure 7

Example Student Response to a Graph Item

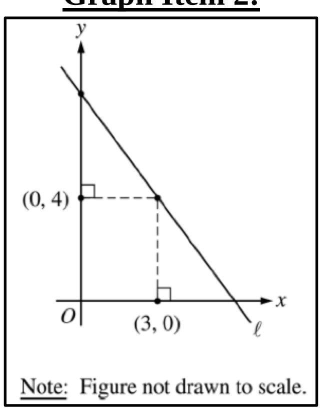
Graph Item 1:  **(G4) What is the equation of line ℓ in the figure above?**	**Student Response:** **We can find our slope by doing rise over run. Here we can see that the line crosses at 2 on the x axis and 3 on the y. So our rise is 3 and our run is 2.** **Next we divide 3 by 2 to get our slope. Therefore, our slope is 3/2.**

Graph Item 2 Student Response 1 (Figure 8) suggests that the student is estimating the point, based on the given information, at which the line crosses the y axis which defines the y-intercept. It suggests that their understanding of graphs is not robust enough to discern the inability to visually identify the y-intercept in this graph similar to the way that it is typically presented. Similarly, Graph Item 2 Student Response 2 represents common errors in using noncollinear points to solve graph problems. This is common in problems in which collinear points are not explicitly presented and furthermore require the use of conceptual knowledge of graphs to use the given information in the graph to ascertain collinear points in order to solve the problem. Thus, both responses suggest erroneous interpretations of the graph shown due to knowledge issues associated with interpretation and representation of graphs.

Additionally, the work in Student Response 1 demonstrated knowledge of the formation of a linear equation, showing that they are familiar with all of the aspects of the equation and where to plug in the given or extracted (albeit incorrect) information, so this response would not be considered a construction error. However, this is also a demonstration of how relying on procedural knowledge of linear functions is not sufficient for correctly solving these types of problems. Because the work in Student Response 2 does not demonstrate familiarity with the correct procedural knowledge required to solve the problem, it would also be considered a type of construction error.

Figure 8

Example Student Responses to a Graph Item

<table>
<tr>
<td rowspan="2"><u>Graph Item 2:</u>


(G1) If the line ℓ has a slope of -2, what is the y-intercept of ℓ?</td>
<td><u>Student Response 1:</u>

slope = rise/run
y = mx + b
b = y-intercept
8 = -2 (0) + b
b = 8</td>
</tr>
<tr>
<td><u>Student Response 2:</u>

4
The y intercept is given by the point (0,4).</td>
</tr>
</table>

There are many variations of these responses which ultimately demonstrate interpretation aligned with elements of the appropriate knowledge and cognitive processes for particular types of items. The responses shown are representations of many common solutions that suggest erroneous interpretations rooted in weak or flawed knowledge of graphs and tables and their relationships to the process of solving linear functions problems. This is in contrast to erroneous constructions of the responses to these items. In a few cases, erroneous responses to these problems were unclear in conveying their interpretation of the given information and how they arrived at their final answer. These responses were coded as "unclear" along the interpretation dimension, but respectively coded along the construction dimension.

Based on the error coding described above, Table 7 provides a summary of all the error subtypes identified for Table and Graph items.

Table 7

Error Types and Subtypes Across Table and Graph items

Item Type	Table Items	Graph Items
Interpretation Errors	Describes change in the x-values and/or y-values separately, either by providing an expression of the change in only x-values, only y-values, or an amalgamation both (T.I.E1).	Visually estimates and/or provides response using the given information at face-value (G.I.E1).
	Describes change in each pair of corresponding x-values and y-values (T.I.E2).	Identifies a given non-collinear point from the graph as relevant information to solve the problem (G.I.E2).
	Response is incorrect and does not clearly show how they arrived at the final answer (source of error cannot be identified) (T.Unclear).	Response is incorrect and does not clearly show how they arrived at the final answer (source of error cannot be identified) (G.Unclear).
Construction Errors	Demonstrates familiarity with the formation of a linear equation, but incorrectly carries out the steps to construct the final solution (T.C.E1).	Demonstrates familiarity with the formation of a linear equation, but incorrectly carries out the steps to construct the final solution (G.C.E1).
	Response does not show evidence of attempting to represent the relationship between the *x* and *y* values in the canonical form of a linear function (y = m x + b) (T.C.E2).	Response does not show evidence of attempting to represent the relationship between the *x* and *y* values in the canonical form of a linear function (y = m x + b) (G.C.E2).

Note. The denotation in each category coincides with the code representation of the respective error subtype for further reference.

F. Scoring Reliability

To evaluate the reliability of the scoring scheme to support the claim that these classifications represent common errors among these responses, two raters scored the responses for 30% of the respondents (n = 50) with the goal of high agreement between both raters on all dimensions of error categorization. Initially, the responses of three respondents were discussed among the raters as a training session. Then, both raters independently coded 10% of the subset and Cohen's Kappa values were derived for each item which resulted in high reliability (M = .89, SD = .05) (Viera & Garrett, 2005). Raters then discussed and resolved all discrepancies, which revealed coding errors that were predominantly non-systematic misclassifications by the raters, but ultimately did not result in any changes to the established scoring scheme. The raters then coded the remaining responses of the subset which resulted in near perfect agreement between both scorers for all items (M = .98, SD = .02). The few resulting discrepancies in the remaining subset were resolved, and again, did not require any changes to the scoring scheme. These results suggest that the scoring scheme for error categorization is a reliable classification tool for the types of errors observed in LF problems presented as tables and graphs and can be used to discern specific conceptual lapses in LF knowledge.

G. Prevalence of Sub-errors by Item Type

Analyzing the frequencies of the subtypes of each error allows for a more precise examination of the specific mis-understandings students have in these actions. To discern the distribution and prevalence of errors across both items and subjects, the frequencies of the error subtypes in each item type group were tabulated. Responses were coded as correct when there was no error and responses with errors were classified based on the specific error observed in the interpretation action and construction action. Table 7 lays out the descriptions of the specific errors in both actions for graph and table problems. Two types of interpretation errors in responses to graph problems were (1) visually estimating and/or using the given information at face-value (G.I.E1) and (2) identifying a given non-collinear point from the graph as relevant information to solve the problem (G.I.E2). When the response is incorrect but does not clearly show how they arrived at the final answer and the source of the error cannot be identified, it was coded as unclear (G.Unclear). Two types of construction errors in graph problems were (1) demonstrating familiarity with the formation of a linear equation, but incorrectly carrying out the steps to construct the final solution (G.C.E1) and (2) not showing evidence of attempting to represent the relationship between the x and y values in the canonical form of a linear function ($y = m x + b$) (G.C.E2).

It is possible for responses to demonstrate different combinations of the types of interpretation and construction errors. It is also possible for responses to demonstrate an error in one action but not in another (e.g., an interpretation error but not a construction error; a construction error but not an interpretation error). The error combinations comprise the specific subtype of interpretation and/or construction errors (and/or no errors) within each item type.

Table 8 displays the frequency of each sub-error combination that was observed in the 982 total responses to the graph problems that were scored. As shown in Table 8, most responses showed no error in the construction actions in (N = 746 - 76%). In contrast, almost half of all 982 responses contained errors in the interpretation actions (N = 459 - 46.7%), and almost two thirds of those interpretation errors demonstrated the G.I.E1 type interpretation error (63.6% -- 292/459). Overall, the single most prevalent error pattern in graph problems was an interpretation error and no construction error with this pattern accounting for 55.1% of all incorrect responses – 290 of 526. Specifically, when solving graph problems students tend to make errors in the interpretation action in which they visually estimate relevant elements of the linear function from the given information and/or use the given information at face-value to solve the problem of generating elements of the equation, and they then correctly carry out the construction actions based on the flawed interpretation. An example of this error is represented in Student Response 1 in Figure 8. Interpretation and construction errors co-occurred in 32.1% of all incorrect responses 169 of 526. Construction errors in the absence of an interpretation error were very infrequent representing 8.6% of all incorrect responses – 45 of 526.

Table 8

Frequency Counts of Error Types on Graph Problems

All Graph Item Responses	**N (observations)**	**NA (missing)**	**Accuracy**			
	982	75	44%			
Error Subtype Frequency		**G.I.E1**	**G.I.E2**	**G.Unclear**	**no Interp. error**	*Marginals*
	G.C.E1	55	39	0	37	*131*
	G.C.E2	35	40	22	8	*105*
	no Const. error	202*	88	0	456	*746*
	Marginals	*292*	*167*	*22*	*501*	*982*

Note. Error frequencies are categorized based on representative error codes (e.g. G.I.E1 refers to a response to a graph problem (G) in which a type 1 (E1) interpretation (I) error was observed. The asterisk (*) denotes the single most frequent error pattern.

Table 7 also lays out the descriptions of the specific errors in interpretation and construction actions for table problems. Two types of interpretation errors in responses to table problems were (1) describing change in the x-values and/or y-values separately, either by providing an expression of the change in only x-values, only y-values, or an amalgamation of both (T.I.E1), and (2) describing change in each pair of corresponding x-values and y-values

(T.I.E2). When the response is incorrect but does not clearly show how they arrived at the final answer and the source of the error cannot be identified, it was coded as unclear (T.Unclear).

Two types of construction errors in table problems were (1) demonstrating familiarity with the formation of a linear equation, but incorrectly carrying out the steps to construct the final solution (T.C.E1), and (2) not showing evidence of attempting to represent the relationship between the *x* and *y* values in the canonical form of a linear function (y = m x + b) (T.C.E2). It is also possible for responses to demonstrate different combinations of the types of interpretation and construction errors.

Table 9 displays the frequency of each sub-error combination that was observed in the 978 total responses to the table problems that were scored. As shown in Table 9, there were 398 total errors and almost all of those contained errors in the construction actions (N = 392 – 98.4%), and most of those errors demonstrated the T.C.E2 type construction error (304/392 – 77.6%). Almost all of the total errors also contained errors in the interpretation actions (351 – 88.2%), and most of those errors demonstrated the T.I.E1 type interpretation error (225/351 – 64.1%). If an interpretation error occurred (351/978 – 35.9%), it was almost always associated with a construction error (345 of 351 cases – 98.3%). Few construction errors occurred without an associated interpretation error (26 of 392 – 6.6%).

As noted above, the most prevalent error combination in table problems was both an interpretation error and construction error. Specifically, when solving table problems students tend to make errors in the interpretation action in which they describe the change in the x-values and/or y-values separately, as well as errors in the construction action in which they don't attempt to represent the relationship between the *x* and *y* values in the canonical form. An example of this combination of errors is represented in the Student Responses shown in Figure 6.

Table 9

Frequency Counts of Error Types on Table Problems

All Table Item Responses	N (observations)	NA (missing)	Accuracy			
	978	79	55%			
Error Subtype Frequency		**T.I.E1**	**T.I.E2**	**T.Unclear**	**no Interp error**	*Marginals*
	T.C.E1	36	31	0	21	*88*
	T.C.E2	183*	95	21	5	*304*
	no Const. error	6	0	0	580	*586*
	Marginals	*225*	*126*	*21*	*606*	*978*

Note. Error frequencies are categorized based on representative error codes (e.g. T.I.E1 refers to a response to a table problem (T) in which a type 1 (E1) interpretation (I) error was observed. The asterisk (*) denotes the single most frequent error pattern.

Across all table and graph problems only a few responses were considered unclear (Graph n = 22; Table n = 21) (Table 8 and 9). Responses that were considered unclear along the interpretation dimension aligned with the G/T.C.E2 classification, in which there was no evidence of attempting to represent the given information in the canonical form of a linear function. This means that responses in which the interpretation of the given information was unclear also did not show familiarity with the elements of a LF equation. Missing responses were also tabulated for each item (Table 8 and 9) to determine if an item was more likely to be skipped. Across all responses to all table and graph problems not many were left blank (Graph n = 79; Table n = 75).

H. <u>Sub-error consistency across subjects and items</u>

The sub-error patterns were analyzed for consistency to determine two things: (1) whether the same types of interpretation and construction errors are exhibited across students for the items within a given problem type, and (2) whether students are internally consistent in the types of interpretation and construction errors they make across items within each given problem type.

1. <u>Consistency across items</u>

Intraclass correlation coefficients (ICC) were used as a reliability measure of sub-error pattern consistency across items within each item type. This measure conveys the extent to which students' responses to these items exhibited the same error patterns across each item type. Two ICC tests were conducted, a test of consistency and a test of absolute agreement. The test of consistency determines whether the error patterns are consistent across items, and the test of absolute agreement determines whether the variation in the individual error subtypes are the same across items. It is useful to obtain both measurements because, although items could be highly correlated in their observed error patterns, it is important that the error subtypes themselves exhibit the same patterns across items. This ensures that all items are designed in such a way that students will solve them in a similar manner, verifying that the item set is a reliable tool to specifically explore students' solution processes, and ultimately, their understanding of the LF domain as manifest by that item type. To do this, the counts of the error subtypes for each item were converted to percentages of the sum of the total errors and entered into the analysis. All were included except for those in the "unclear" category. ICC values indicate the level of reliability observed: a value of 0.5 suggests poor reliability, a value between 0.5 and 0.75 suggests moderate reliability, a value between 0.75 and 0.9 suggests good reliability, and values greater than 0.90 suggests excellent reliability (Koo & Li, 2016). The consistency and absolute agreement tests both resulted in moderate to excellent ICC values for the table items (Table 10) and graph items (Table 11). This means that the error patterns varied consistently across items, and that the individual subtype patterns did not vary much from item to item. This indicates that all responses within each item type showed consistent error patterns. Thus, all items are consistent in their ability to uncover how students carry out interpretation and construction actions in these types of LF problems and that they are all designed to call upon the same knowledge.

Table 10

ICC Values and Confidence Intervals for the Test of Consistency and the Test of Absolute Agreement for Table Items

	ICC	**95% CI**
Consistency	0.86	0.64 - 0.96
Agreement	0.82	0.54 - 0.95

Table 11

ICC Values and Confidence Intervals for the Test of Consistency and the Test of Absolute Agreement for Graph Items

	ICC	**95% CI**
Consistency	0.85	0.62 - 0.96
Agreement	0.84	0.59 - 0.96

2. Consistency across subjects.

It is useful to analyze the error consistency on the subject level because it will ensure that individual students are consistently calling upon the same knowledge to solve each of these types of problems. This will indicate whether the knowledge that is called upon is a reflection of the students' firm understanding, or misunderstanding, of the domain as manifest within that problem type. To test whether subjects are internally consistent in the types of errors they make, a subset of subjects were evaluated on the basis of the consistency of their error patterns across items. A binomial probability calculation was done to identify the subset as well as the criteria for determining the reliability of students' error patterns. Because there is a 33% chance that an erroneous response could contain any of the error types (interpretation error, construction error, or both), a subject would be considered reliable if the probability of their error patterns are significantly beyond the 33% chance of random distribution between the error types. The calculation suggested that subjects who made errors in at least 5 of the 7 items in each item type should be analyzed for a stable estimate of reliability and consistency. In graph problems, 95 (62.9%) subjects met this criterion, and in table problems, 89 (58.9%) subjects met this criterion. The error patterns of these subjects were each evaluated based on criteria that signal that the error patterns observed have less than 5% chance of occurring at random. Based on the binomial probability calculation, the criteria are as follows: If they made errors on 5 items, at least 4 must be the same error type ($p = 0.03$); if they made errors on 6 items, at least 5 must be the same error type ($p = 0.02$); if they made errors on 7 items, at least 6 must be the same error type ($p = 0.01$). Subjects' total numbers of each of the three error types were evaluated based on these criteria. Only those who met the criteria were considered reliable. Of the 95 subjects in the subset for graph problems, 70.5% were reliable. Of the 89 subjects in the subset for table problems, 75.3% of subjects were reliable. This suggests that a large portion of subjects who made errors on more than half of all items in each item type were consistent in their error patterns. This supports the conclusion that as students solve each of these types of problems, they are consistently calling upon the same knowledge and verifies that their responses are a true reflection of their understanding, or misunderstanding, of the domain as manifest within that problem type.

I. Sub-error patterns in high and low performers

It is useful to compare the action-level error patterns in these problems for high and low performers. This will highlight the respective discrepancies in the schemas of those who are better at solving these problems compared to those who are worse. To do this, the highest performing 27% (n = 39) and lowest performing 27% (n = 39) of subjects (Kelley, 1939) were extracted and compared on the basis of their error patterns on table problems and on graph problems independently. Tables 12 and 14 show the frequencies of each error subtype in graph items for top and bottom performance groups, respectively. Tables 13 and 15 show the frequencies of each error subtype in table items for top and bottom performance groups, respectively.

As shown in Table 12, out of the total 272 responses of the top performing groups on graph items, 52 (19.1%) contained errors. Of those erroneous responses, most were construction errors that are not associated with interpretation errors (34/52 - 65.4%), and most of those demonstrated the G.C.E1 type construction error (25/34 - 73.5%). As shown in Table 13, out of the total 274 responses of the top performing groups on table items, 39 (14.2%) contained errors. Of those erroneous responses, most were construction errors that are not associated with interpretation errors (27/39 - 69.2%), and most of those demonstrated the G.C.E1 type construction error (21/27 - 77.8%). Both groups predominantly make only construction errors, and specifically one in which they are demonstrating familiarity with the formation of a linear equation, but incorrectly carrying out the steps to construct the final solution.

Table 12

Frequency Counts of Error Types on Graph Problems for the Top Performing 27% of Subjects

	G.I.E1	**G.I.E2**	**Correct (no error)**	*Totals*
G.C.E1	**2**	**0**	**25***	27
G.C.E2	**0**	**1**	**9**	10
Correct (no error)	**11**	**4**	**220**	235
Totals	13	5	254	272

Note. The asterisk (*) denotes the error subtype that occurred the most in each performance *group.*

Table 13

Frequency Counts of Error Types on Table Problems for the Top Performing 27% of Subjects

	T.I.E1	**T.I.E2**	**Correct (no error)**	*Totals*
T.C.E1	**5**	**0**	**21***	26
T.C.E2	**2**	**1**	**6**	9
Correct (no error)	**3**	**1**	**235**	239
Totals	10	2	262	274

Note. The asterisk (*) denotes the error subtype that occurred the most in each performance *group.*

As shown in Table 14, out of the total 259 responses of the bottom performing groups on graph items, 221 (85.3%) contained errors. Of those erroneous responses, most were interpretation errors (218/221 - 98.6%), and the single most prevalent error is one that demonstrated the G.I.E1 type interpretation error with no associated construction error (89/218 - 40.8%). So, when solving graph problems low performing students tend to make errors in the interpretation action in which they visually estimate relevant elements of the linear function from the given information and/or use the given information at face-value to solve the problem of generating elements of the equation, and then correctly carry out the construction actions based on the flawed interpretation.

Table 14

Frequency Counts of Error Types on Graph Problems for the Bottom Performing 27% of Subjects

	G.I.E1	**G.I.E2**	**Correct (no error)**	*Totals*
G.C.E1	**45**	**20**	**2**	67
G.C.E2	**10**	**25**	**1**	36

Correct (no error)	**89***	**29**	**38**	156
Totals	144	74	41	259

Note. The asterisk (*) denotes the error subtype that occurred the most in each performance *group.*

As shown in Table 15, out of the total 261 responses of the bottom performing groups on table items, 241 contained errors (92.3%). Of those erroneous responses, most were those that contained both interpretation and construction errors (227/241 - 94.2%). The most prevalent error in those responses were those that demonstrated both the T.I.E1 type interpretation error and T.C.E2 type construction error (121/227 - 53.3%). So, when solving table problems low performing students tend to make errors in the interpretation action in which they describe the change in the x-values and/or y-values separately, as well as errors in the construction action in which they don't attempt to represent the relationship between the *x* and *y* values in the canonical form.

Table 15

Frequency Counts of Error Types on Table Problems for the Bottom Performing 27% of Subjects

	T.I.E1	**T.I.E2**	**Correct (no error)**	*Totals*
T.C.E1	**24**	**19**	**12**	55
T.C.E2	**121***	**63**	**0**	184
Correct (no error)	**2**	**0**	**20**	22
Totals	147	82	32	261

Note. The asterisk (*) denotes the error subtype that occurred the most in each performance *group.*

Error patterns observed in the Bottom performance groups are consistent with the error patterns observed across all subjects as a whole in both graph and table items, while the Top performance groups seemed to predominantly make construction errors in both graph and table items. These findings reveal that schemas in those who are better at solving these problems may only exhibit discrepancies in the procedural aspect of their LF knowledge. However, those who are worse at solving these problems exhibit discrepancies in aspects of the problem-solving process that may reflect fundamental conceptual misunderstandings in their LF knowledge, similar to what was found with the larger group (across all subjects).

These error analyses complement the psychometric reliability analyses in that there are highly consistent errors among each problem type, which is consistent with the high internal consistency and discriminability which indicate that students are behaving similarly with all the items in each subset. Even with high difficulty levels of the items, students are consistent in their item-level performance across each subset, including the specific mistake they make in the interpretation and construction actions of the solution process. In other words, when students make a mistake on one item type, they are likely to make mistakes on many other items of that same item type, and the mistakes they do make are consistent across all items in the subset.

J. <u>Demographic predictors of performance</u>

Students' demographic information was collected based on a self-report questionnaire that includes their major, scores on the ACT/SAT, math and physics classes taken in college thus far and the average grade earned. To determine the demographic variables that are the best predictors of overall performance on these LF items, a backward stepwise fitting of a series of regression models was conducted to explore the individual and additive effects of these variables. The models were evaluated to determine (1) the most parsimonious model that would best fit the data, and (2) the variables that account for the most variance in student performance on LF items. To do this, subsequent changes in the R^2 values of the models were evaluated for significant reductions. Only 20% of subjects in the sample had taken physics courses and only

28% reported ACT scores, so both variables were not included in the models. SAT scores were transformed to z-scores and average grades in math classes and major were manually coded to represent the levels of each.

Table 16 shows 6 models for which Models 1, 2, 3 include all predictors with 1 predictor subsequently removed, and Models 4, 5, 6 include each predictor on its own. Model 1, which includes SAT scores, major, and math grades, $R^2 = .60$, was compared to Model 2, which includes SAT and major, $R^2 = .59$. The difference was not significant, $F(3, 142) = 1.30$, $p = 0.28$, so Model 2 is a better fit as it is more parsimonious than Model 1. Model 2 was compared to Model 3, which includes SAT and math grade, $R^2 = .57$. The reduction in R^2 between the models was significant, $F(1, 145) = 6.83$, $p < 0.001$, so Model 2 is a better fit. Models 4, 5, and 6 each independently comprise math grade, major, and SAT, respectively. Model 2 was compared to each model to determine if these variables alone can predict performance. All three models resulted in a significant reduction in R^2 compared to Model 2, such that math grade, major, and SAT alone are not good predictors of performance compared to the combined effects of SAT and major [Model 4, R^2 =.05: $F(2, 145) = 97.57$, $p < 0.001$; Model 5, R^2 =.04: $F(1, 145) = 197.33$, $p < 0.001$; Model 6, R^2 =.57: $F(4, 145) = 2.52$, $p < 0.05$]. Thus, across all models the most significant and parsimonious model is Model 2. This means that, while math course performance does add a bit to the level of overall prediction when added to SAT, which accounted for the largest amount of the variance (as seen in Model 6), the combined effects of SAT scores and major, specifically a major in engineering, seem to be strong predictors of overall performance on all the LF items, $R^2 = .59$, $F(5, 145) = 42.37$, MSE = 2.53, $p < 0.01$. However, although Model 2 is best, SAT has accounted for a substantial portion of the variance. So, while Model 2 is the best fit, not much predictive ability is lost with using only SAT scores without major. This means that students' performance on the SAT, and possibly their major, is a reasonable index of the math-related prior knowledge that they possess, and serve as potential indicators of their understanding of linear functions. Further exploration of this effect could be pursued to understand how these measures relate to students' performance on LF problems. Exploring this could also inform the process of identifying students who could potentially benefit from an intervention to improve their LF knowledge.

Table 16

Models for Which the Data was Fitted that Differ Based on the Set of Predictors in a Backward Fitting of a Series of Regressions

	Dependent variable:					
	Overall.Score					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	5.73^{***} (1.10)	5.49^{***} (1.03)	8.20^{***} (0.38)	8.94^{***} (0.56)	5.50^{***} (1.58)	7.74^{***} (0.21)
SAT	3.10^{***} (0.23)	3.16^{***} (0.22)	3.07^{***} (0.23)			3.13^{***} (0.23)
MAJOREngineering	2.93^{***} (1.09)	2.68^{**} (1.07)			2.39 (1.64)	
MAJORLiberal.Arts	-0.03 (2.79)	-0.70 (2.76)			-5.50 (4.19)	
MAJORMath	2.33^{*} (1.33)	2.07 (1.31)			2.50 (2.00)	
MAJORScience	2.05^{*} (1.13)	1.81 (1.10)			2.36 (1.69)	
MATH.GRADEB	-0.66 (0.48)		-0.58 (0.49)	-1.67^{**} (0.72)		
MATH.GRADEC	-1.00 (0.62)		-1.01 (0.63)	-2.08^{**} (0.93)		
MATH.GRADED	0.88 (1.54)		0.36 (1.54)	-0.27 (2.30)		
Observations	151	151	151	151	151	151
R^2	0.60	0.59	0.57	0.05	0.04	0.57
Adjusted R^2	0.58	0.58	0.56	0.03	0.01	0.56
Residual Std. Error	2.53 (df = 142)	2.53 (df = 145)	2.58 (df = 146)	3.86 (df = 147)	3.88 (df = 146)	2.59 (df = 149)
F Statistic	27.14^{***} (df = 8; 142)	42.37^{***} (df = 5; 145)	49.29^{***} (df = 4; 146)	2.41^{*} (df = 3; 147)	1.55 (df = 4; 146)	193.90^{***} (df = 1; 149)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

IV. GENERAL DISCUSSION

Multiple quantitative and qualitative analyses of student responses to three different types of LF items provided valuable information about college students' ability to solve such problems and, more specifically, the types of errors they make and the extent of their struggles with materials representing this important subdomain of mathematics. Analyses showed that each of the item types can distinguish between high and low performing subjects based on overall accuracy within each problem subset and across the entire set of items, as well as specific error patterns that high and low performers tend to make on the two most challenging of the problem types – tables and graphs.

Overall, results showed that students performed better on Description items compared to Table and Graph items, consistent with prior data (Mielicki et al., 2019). Relatively high accuracy on Description items suggests that, generally, students possess the knowledge and ability to carry out interpretation and construction actions necessary to solve problems in which the information is verbally presented and a simple linear equation is the desired product. An instructional intervention that focuses on improving the knowledge and relevant cognitive processes associated with solving Description items is not necessarily warranted. This is in contrast to the relatively low accuracy observed in solutions to Table and Graph items and the respective misunderstandings identified from error analysis for both problem types. The results suggest erroneous interpretations of the given information in such problems and erroneous constructions of the responses to these items. It is argued that such errors are rooted in weak or flawed conceptual knowledge of graphs and tables with respect to their use in the process of solving linear functions problems. The error analysis suggests a need to address the specific knowledge discrepancies, conceptual and procedural, that have been identified for both problem types.

The present findings can be interpreted using two complementary frameworks: (1) a mathematical knowledge framework and (2) a cognitive processing framework. To successfully solve a linear functions problem, one must have specific forms of relevant linear functions knowledge and understanding which are in turn required to execute key aspects of a cognitive problem-solving process for specific types of LF problems. For example, cognitive theory suggests that one must be able to build a proper representation of the situation presented in the problem which requires the execution of cognitive processes that call upon relevant domain knowledge (e.g., Kintsch & Greeno, 1985). If one lacks relevant conceptual knowledge, or if it is incorrect, the situation representation will demonstrate a lack of understanding that will then carry forward into the problem solution. Analyzing the traces of solutions to linear functions problems can provide insight into students' competency levels in a domain, specifically, the understanding they have and their ability to map that to a representation of the problem.

In the solution of the LF problems studied in this project, students' solutions can be interpreted with respect to their knowledge and understanding relative to articulations of the same in the Common Core standards for mathematics. These standards speak to the type of knowledge one must have to be able to carry out each action in the solution process and what evidence of that looks like when executing each activity of the overall solution process. Specifically, we are interested in the knowledge that gets called upon and used during interpretation of the given problem information and construction of the final solution. While it is crucial to identify what knowledge is and is not available, it is also important to understand how this knowledge gets called upon in executing each action. This can be explored by looking into the cognitive underpinnings of the solution process which is conveyed by the students' schematic knowledge for representation of the given problem. Being able to correctly recognize and classify problem types allows for information given in the problem to be represented correctly and efficiently which then allows for the solution procedure to be carried out correctly.

A. <u>Mathematical Implications of Rational Errors</u>

Student responses demonstrate some disparities between what they know and what they should know to be able to solve LF graph and table problems with respect to the Common Core standards. For graph problems students failed to “Interpret key features of graphs...for a function that models a relationship between two quantities (Standard F-IF.4)”, but they did show familiarity with the steps to “Construct linear...functions...given a graph (Standard F-LE.2)” (CCSSI, 2010). For table problems students did not demonstrate the understanding that “...a function from one set (called the domain) to another set (called the range) assigns to each element of the domain exactly one element of the range (Standard F-IF.1)”, and this led to failing to “Construct linear...functions...given...two input-output pairs (including reading these from a table) (Standard F-LE.2)” (CCSSI, 2010).

The most pervasive observation is that the students understand the simple linear equation from a procedural perspective and their conceptual or schematic understanding of what this equation represents seems to be lacking. For example, while they do understand and are familiar with the linear equation, they lack the schematic representation of a table that actually relates the changes in *x* relative to the changes in *y* as a functional relationship, and instead they treat them as individual pairs. Because of this, they are not able to build a proper representation of the situation presented in the problem, and thus, cannot correctly carry out the solution procedure.

B. <u>Cognitive Implications of Rational Errors</u>

There is evidence that suggests that the errors observed are not random and illogical and do not stem from general unfamiliarity with the LF domain as a whole. High consistency in the error patterns suggests that students are overwhelmingly and consistently making the same errors across items of a given type. One error is that students recognized the relevant approach to the graph problems by applying their knowledge of LF, evidenced by showing familiarity with the elements of the canonical form, but they did not recognize this relevance with table problems. Students appear to make surface-level interpretations about the given information based on the structure in which it is presented. This means that the knowledge regarding the procedural approach for linear functions problems was effectively called upon for graph problems but not for table problems. For example, while solving table problems, students do not call upon the relevant linear function knowledge, even if they do possess it. This is because, regardless of other cues in the problem, when seeing a problem in which information is structured in a table format with discrete values, it is not classified as a LF problem, and therefore their LF knowledge is not called upon. This suggests that when these students are solving this type of problem, the knowledge that is called upon is not relevant to this problem type which leads to a flawed representation of the situation. Their approach to solving the problems, however, was logically consistent and systematic, but was not the correct approach for the goal of the problem. This shows that the errors observed are rational errors and are not random and/or due to a fundamental lack of knowledge of LF. Students are in fact familiar with the domain but seem to make systematic errors in the interpretation and construction actions that suggest that they lack the ability to properly build representations of these problems.

The combination of errors that are most prevalent in table and graph problems can be presumed to be indicative of the cognitive underpinnings that lead to the apparent misunderstandings. A taxonomy for rational errors developed by Ben-Zeev (1998) instantiates the mathematical reasoning processes which reveal underlying cognitive mechanisms. The basis of this framework is that, as observed in the results of this study, errors in problem solving are often systematic and logically consistent rather than random and unsubstantiated (Ben-Zeev, 1995; 1996). This rational error framework maps out what is referred to as *inductive failures* that seem to arise when one overgeneralizes relevant conceptual and procedural knowledge from prototypical examples which result in erroneous solutions to math problems. One type of inductive failure is the process of overgeneralizing a solution process based on familiar structural features, referred to as *syntactic induction*. An example of syntactic induction is a *misspecification of constraints* (Ben-Zeev, 1998), which is the tendency to adapt a relevant known rule to the problem without appropriate constraints (Matz, 1982). The predominant erroneous solutions observed in responses to the graph problems align with this kind of error such that responses showed only interpretation errors. While correct construction actions showed evidence of familiarity with the aspects of a linear equation and where to plug in extracted (albeit incorrect) information, the erroneous interpretation action suggests a discrepancy in the understanding of how the elements of a function describe the properties of a line. This is evidenced by the lack of constraint specification when the information was extracted by estimation or at face-value and greatly relying on procedural knowledge, rather than applying the relevant rule which comprises the conceptual mapping of all information needed to be extracted from the graph to the elements of the equation needed to solve the problem. This suggests that they are interpreting the information in the graph by only using its surface-structural features, reflecting a flawed representation of the problem.

Common errors observed in responses to table problems also seem to stem from the misspecification of constraints. The solution to this problem type requires systematic evaluation of all given values and inferring the pattern of relations of the values relative to each other. However, students appear to have a biased conceptual understanding of linear functions in the context of discerning and expressing given value patterns in which they show a lack of consideration of the constraints during the process of formulating an equation that represents the pattern of the given values. Specifically, responses showed that the patterns of the *x* and *y* values were evaluated separately, disregarding the rule which requires the evaluation of multiple corresponding *x* and *y* values to execute the appropriate pattern induction process required to solve these problems. Consequently, this is reflected in their construction actions such that their conceptual understanding of pattern induction does not align with the subsequent procedure which comprises the process of solving for the elements of the LF equation then constructing it. Instead, responses were constructed in a way that aligns with their biased understanding of pattern induction and extrapolation. The errors observed in both graph and table problems demonstrate the common tendency for students to think of tables and graphs as just a compilation of individual points, instead of understanding that those points make up a conceptual entity (Schoenfeld et al., 1993; Stein et al., 1990; Yerushalmy, 1988). Thus, it seems that the surface-structural characteristics of these problems contribute to the erroneous solutions of these types of LF problems.

C. <u>Implications for Instructional Interventions</u>

The present research focused on identifying and modeling errors that students most commonly make during the essential interpretation and construction actions of the problem-solving process for different types of LF items. The errors reflected common misunderstandings that students have in this domain for specific classes of problems. It revealed that students tend to make rational rule-based errors that are conditional on the format in which the problem information is presented. One goal of the research was to provide information for an instructional intervention that attends to the specific deficiencies that students have in their knowledge of LF and explores the learning benefits of different intervention designs, specifically, learning from worked examples with and without error detection/correction. There is considerable research showing that learning from worked out examples in general promotes deep understanding during the learning process because it maximizes germane cognitive load (Große & Renkl, 2007). In particular, students can allot more cognitive resources to the learning of important content and less to the demands of performance expectations (Renkl et al., 2003). Research also shows evidence of learning benefits when students are presented with both correct and incorrect information (Große & Renkl, 2007; Schnotz, 2001; Schnotz, Vosniadou, & Carretero, 1999) in such a way that it triggers impasse-driven episodes (VanLehn, 1999). According to VanLehn (1999), learning outcomes can be promoted through experiencing impasses, which prompt learning episodes. These impasses are characterized by noticing and reflecting on errors and are critical to knowledge acquisition. A study under development aims to trigger what VanLehn (1999) refers to as impasse-driven episodes during both error detection and error correction conditions to facilitate learning outcomes and performance on subsequent problems.

In general, learning from error detection and correction can ameliorate the struggles that students experience because they can appropriately implement procedural steps to solve the problems while also understanding why these procedures work (Gott, 1993), which according to Catrambone (1996, 1998), and Gott et al. (1993), is crucial to be able to flexibly apply pertinent aspects of learned procedures while solving problems. Having determined the most predominant errors that students make allows for the intervention to both address specific inductive failures and strategically trigger an impasse by presenting them in worked out examples. However, there are some factors that could affect the efficacy of the interventions and how beneficial they could be to students. For example, prior knowledge of the domain mediates the efficacy of learning from error detection and correction because lack of prior knowledge can cause extraneous cognitive load (Paas & Van Gog, 2006). Additionally, the findings from this study showed that ACT/SAT scores can be predictive of students' performance on these types of problems, so it could be useful to explore this factor as a criterion for determining the student demographic that would most benefit from an intervention.

Learning from worked examples with error detection and correction is a potentially favorable approach given the nature of the errors observed in the responses to these problems which showed that students tend to rely on their knowledge of the concept's 'prototypical examples' instead of the definition of that concept as seen in the biases demonstrated in the student's responses. This also aids in thinking analytically and promoting adaptive conceptual thinking which helps move students away from the tendency to overly rely on procedural knowledge (Dorier & Sierpinska, 2001) similar to the misspecification of constraints observed in the responses in this study.

Appendix A

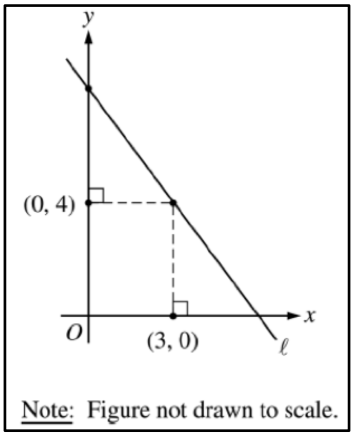
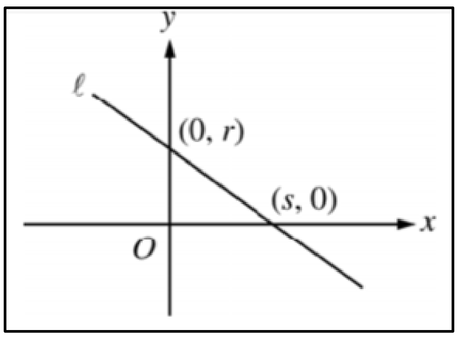
Item	p - value
A plumber charges customers $48 for each hour worked plus an additional $9 for travel. If h represents the number of hours worked, write an expression that could be used to calculate the plumber's total charge in dollars?	0.86
If m represents the total number of months that Jill worked and p represents Jill’s average monthly pay, write an expression that represents Jill’s total pay for the months she worked.	0.83
A car can seat c adults. A van can seat 4 more than twice as many adults as a car can. Write an expression in terms of c that describes how many adults the van can seat.	0.91
The admission price to a movie theater is $7.50 for each adult and $4.75 for each child. Write an equation to determine T, the total admission price, in dollars, for x adults and y children.	0.89
If □ represents the number of newspapers that Lee delivers each day, write an expression that represents the total number of newspapers that Lee delivers in 5 days.	0.93
The cost, C, of printing greeting cards consists of a fixed charge of 100 cents and a charge of 6 cents for each card printed. Write an equation that can be used to determine the cost of printing n cards.	0.92
A taxi company had a basic charge of 25 zeds and a charge of 0.2 zeds for each kilometer that the taxi is driven. Write an expression that represents the cost in zeds to hire a taxi for a trip of n kilometers.	0.88

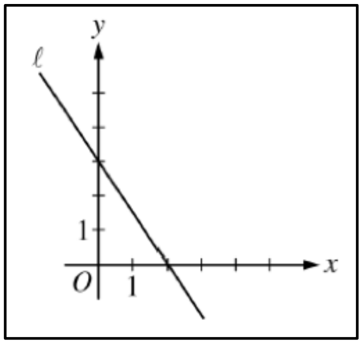
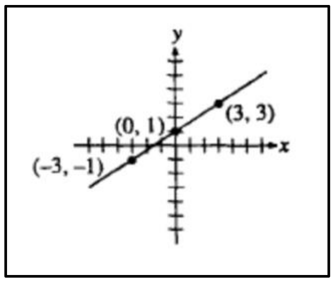
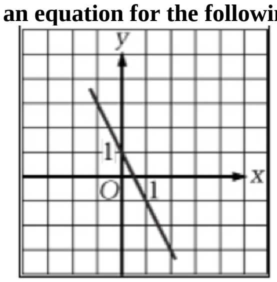
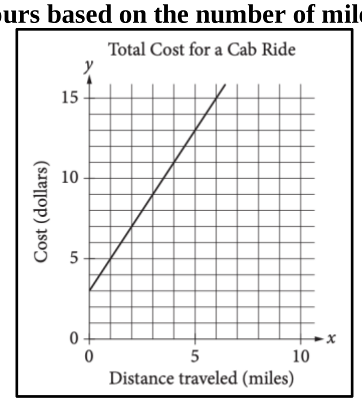
Appendix B

<table>
<tr><th>Item</th><th>p - value</th></tr>
<tr><td>
This table represents a relation between x and y.
Write an equation that could represent the relation.
<table>
<tr><th>x</th><th>y</th></tr>
<tr><td>1</td><td>1</td></tr>
<tr><td>2</td><td>4</td></tr>
<tr><td>3</td><td>7</td></tr>
<tr><td>4</td><td>10</td></tr>
</table>
</td><td>0.83</td></tr>
<tr><td>
<table>
<tr><th>x</th><th>y</th></tr>
<tr><td>0</td><td>-3</td></tr>
<tr><td>1</td><td>-1</td></tr>
<tr><td>2</td><td>1</td></tr>
</table>
Write an equation which describes the three pairs of x and y in this table.
</td><td>0.82</td></tr>
<tr><td>
<table>
<tr><th>x</th><th>y</th></tr>
<tr><td>1</td><td>7.5</td></tr>
<tr><td>2</td><td>13.0</td></tr>
<tr><td>3</td><td>18.5</td></tr>
<tr><td>4</td><td>24.0</td></tr>
</table>
Write an equation that expresses y in terms of x for each of the four pairs of values shown in the table above?
</td><td>0.75</td></tr>
<tr><td>
<table>
<tr><th>x</th><th>y</th></tr>
<tr><td>0</td><td>-1</td></tr>
<tr><td>1</td><td>2</td></tr>
<tr><td>2</td><td>5</td></tr>
<tr><td>3</td><td>8</td></tr>
<tr><td>10</td><td>29</td></tr>
</table>
Write an equation that represents the relationship between x
</td><td>0.77</td></tr>
</table>

<table>
<tr><td>and y shown in the table above?</td><td></td></tr>
<tr><td>The table below lists coordinates of several points on a line. Write an equation of the line.
<table><tr><th>x</th><th>y</th></tr><tr><td>1</td><td>3</td></tr><tr><td>2</td><td>5</td></tr><tr><td>3</td><td>7</td></tr><tr><td>4</td><td>9</td></tr></table></td><td>0.86</td></tr>
<tr><td>As part of a lesson on motion, students observed a cart rolling at a constant rate along a straight line. As shown in the chart below, they recorded the distance y feet, of the cart from a reference point at 1 - second intervals from t = 0 seconds to t = 5 seconds.
<table><tr><td>t</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>y</td><td>14</td><td>19</td><td>24</td><td>29</td><td>34</td><td>39</td></tr></table>Write an equation that represents the data.</td><td>0.77</td></tr>
<tr><td>The table shows a relation between x and y.
<table><tr><td>x</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>y</td><td>7</td><td>10</td><td>13</td><td>16</td></tr></table>Write an equation that expresses this relation.</td><td>0.86</td></tr>
</table>

Appendix C

Item	*p* - value
 **If the line ℓ has a slope of -2, what is the y-intercept of ℓ?**	**0.21**
 **In the figure above, what is the slope of the line ℓ?**	**0.72**
 **The graph above shows the number of George's unsold candy bars over a 10-day period. Write an equation for the line that all the points on the graph lie.**	**0.22**

<table>
<tr><td>
What is the equation of line ℓ in the figure above?</td><td>0.67</td></tr>
<tr><td>
What is the slope of the line shown in the graph above?</td><td>-</td></tr>
<tr><td>Provide an equation for the following graph.
</td><td>0.57</td></tr>
<tr><td>The line graphed in the xy-plane below models the total cost, in dollars, for a cab ride, y, in a certain city during non-peak hours based on the number of miles traveled, x

According to the graph, what is the cost for each additional mile traveled, in dollars, of a cab ride?</td><td>-</td></tr>
</table>

Appendix D

1. Current **age** (in years): ______
2. **Gender**:
    - Female
    - Male
    - Other
3. Highest level education currently **completed**:
    - High School / G.E.D
    - Bachelor's
    - Master's
4. Score on **ACT**: ________
5. Score on **SAT**: ________
6. What is your **major**? _________________
7. What **career path** are you working toward? _________________
8. Number of **math** courses taken in college: ___
9. Average **grade earned in math** courses: ___
10. Number of **physics** courses taken in college: ___
11. Average **grade earned in physics** courses: ___
12. Highest level of education completed by **mother**:
    - None
    - High school / G.E.D
    - Professional Training / Certificate
    - Some College
    - College
    - Some Graduate
    - Master's
    - Doctorate (Ph.D. / M.D. / J.D.)
13. Highest level of education completed by **father**:
    - None
    - High school / G.E.D

- Professional Training / Certificate
- Some College
- College
- Some Graduate
- Master's
- Doctorate (Ph.D. / M.D. / J.D.)

REFERENCES

Alves-Dias, M. and Artigue, M. (1995). Articulation Problems Between Different Systems of Symbolic Representations in Linear Algebra. In The Proceedings of PME19, *2* (34-41) Universidade Federal de Pernambuco, Recife, Brazil..

Andersen, E. B. (1973). A goodness of fit test for the Rasch model. Psychometrika, 38, 123- 140.

Ben-Zeev, T. (1995). The nature and origin of rational errors in arithmetic thinking: Induction from examples and prior knowledge. Cognitive Science, 19, 341-376.

Ben-Zeev, T. (1996). When erroneous mathematical thinking is just as "correct": The oxymoron of rational errors. In R. J. Sternberg & T. Ben-Zeev, (Eds.), The nature of mathematical thinking (pp. 55-79). Mahwah, NJ: Erlbaum.

Ben-Zeev, T. (1998). Rational Errors and the Mathematical Mind. Review of General Psychology, 2(4), 366–383.

Britton, S., & Henderson, J. (2009). Linear algebra revisited: An attempt to understand students' conceptual difficulties, *International Journal of Mathematical Education in Science and Technology*, *40*(7), 963-974, DOI: 10.1080/00207390903206114

Catrambone, R. (1996). Generalizing solution procedures learned from examples *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1020-1031.

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems *Journal of Experimental Psychology: General, 12,* 355-376.

Carlson, D. (1997) Teaching linear algebra: Must the fog always roll in? in Resources for Teaching Linear Algebra, Mathematical Association of America, Washington, (39–51).

Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied, 4*(2), 75.

Common Core State Standards Initiative. (2010). Common core state standards for mathematics. Retrieved from http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Conley, D., Drummond, K., de Gonzalez, A., Rooseboom, J., & Stout, O. (2011). Reaching the goal: The applicability and importance of the common core state standards to college and career readiness. Eugene, OR: Educational Policy Improvement Center.

Cummings, D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. Cognitive Psychology, 10, 405-438.

Curry, L. A. (2004). The effects of self-explanations of correct and incorrect solutions on algebra problem-solving performance. In K. Forbus, D. Gentner, & T. Regier (Eds.), Proceedings of the 26th Annual Conference of the Cognitive Science Society (p. 1548). Mahwah, NJ: Erlbaum.

Dorier, J. L. (Ed.). (2000). On the teaching of linear algebra (Vol. 23). Springer Science & Business Media.

Dorier, J-L., Robert, A., Robinet, J., and Rogalski, M. (2000). The obstacle of formalism in linear Algebra, in On the Teaching of Linear Algebra, J-L. Dorier, ed., Kluwer, Dordrecht, Boston, London, pp. 85–124.

Dorier, J. L., & Sierpinska, A. (2001). Research into the teaching and learning of linear algebra. In The teaching and learning of mathematics at university level (pp. 255-273). Springer, Dordrecht.

Gott, S.P., Parker, E., Pokorny, R.A., E. Dibble, Glaser, R. (1993). A naturalistic study of transfer: Adaptive expertise in technical domains D.K. Detterman, R.J. Sternberg (Eds.), Transfer on trial: Intelligence, cognition, and instruction, Ablex, Norwood, NJ, pp. 258-288.

Große, C. S., & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes?. *Learning and instruction*, *17*(6), 612-634.

Hillel, J. (2000). Modes of Description and the Problem of Representation in Linear Algebra. In J-L Dorier (Ed.), On the Teaching of Linear Algebra, pp. 191-207. Dordrecht: Kluwer Academic Publishers.

Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.1. https://CRAN.R-project.org/package=stargazer

Janvier, C. (1987). Translation processes in mathematics education. In C. Janvier (Ed.), Problems of representation in mathematics learning and problem solving (pp. 27-31). Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P. N., & Byrne, R. M. J. (1993). Precis of deduction. Behavioral and Brain Sciences, 16, 323e380.

Kelley, T. L. (1939). The selection of upper and lower groups for the validation of test items. Journal of Educational Psychology, 30(1), 17–24. doi:10.1037/h0057123.

Kintsch, W., & Greeno, J. (1985) Understanding and Solving Word Arithmetic Problems. *Psychological Review, 92*(1), 109-129

Klauer, K. J. (2001). Training des induktiven Denkens. In K. J. Klauer (Ed.), Handbuch kognitives training (2nd revised ed.) (pp. 165e209). Go¨ttingen: Hogrefe.

Klieme, E., & Baumert, J. (2001). TIMSS-Impulse für Schule und Unterricht: Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumentation.

Koo, T. K., & Li, M. Y. (2016). Cracking the code: providing insight into the fundamentals of research and evidence-based practice a guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155-163.

Leinhardt, G., Zaslavsky, O., & Stein, M. (1990). Functions, Graphs, and Graphing: Tasks, Learning, and Teaching. *Review of Educational Research, 60*(1), 1-64. Retrieved February 16, 2020, from www.jstor.org/stable/1170224.

Lohse, J. (1991, March). A cognitive model for the perception and understanding of graphs. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 137-144).

Neubert, G. A., & Binko, J. B. (1992). Inductive reasoning in the secondary classroom. Washington DC: National Education Association.

Markovits, Z., Eylon, B., & Bruckheimer, M. (1986). Functions today and yesterday. *For the Learning of Mathematics, 6*(2), 18-28.

Matz, M. (1982). Towards a process model for high school algebra errors. In D. S lee man & J. S. Brown (Eds.), Intelligent tutoring systems (pp. 25—49). New York: Academic Press.

Mielicki, M. Martinez, M. DiBello, L., Hassan, A-L., & Pellegrino, J. W. (2019). Assessing Mathematics Knowledge & Skill: What College Students Actually Know and Can Do. In O. Zlatkin-Troitschanskaia (Ed.), *Frontiers and Advances in Positive Learning in the Age of InformaTiOn (PLATO)*. New York: Springer.

Paas, F., & Van Gog, T. (2006). Optimising worked example instruction: Different ways to increase germane cognitive load.

Piaget, J., Grize, J., Szeminska, A., & Bang, V. (1968/1977). Epistemology and psychology of functions. Boston: D. Reidel.

Renkl, A., Gruber, H., Weber, S., Lerche, T., & Schweizer, K. (2003). Cognitive Load beim Lernen aus Lo¨sungsbeispielen [Cognitive load when learning from worked examples]. Zeitschrift fu¨r Pa¨dagogische Psychologie, 17, 93e101.

Robert, A. & Robinet, J. (1989). Quelques résultats sur l'apprentissage de l'algèbre linéaire en première année de DEUG, Cahier de Didactique des Mathématiques n°53, IREM de Paris VII.

Schoenfeld, A. S. (1985). Mathematical problem solving. Orlando, FL: Academic Press.

Schoenfeld, A. H. (1988). *When Good Teaching Leads to Bad Results: The Disasters of "Well-Taught" Mathematics Courses. Educational Psychologist, 23(2), 145–166.*

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), Handbook for Research on Mathematics Teaching and Learning (pp. 334-370). New York: MacMillan.

Schoenfeld, A. H., Smith, J. P., & Arcavi, A. (1993). Learning: The microgenetic analysis of one student's evolving understanding of a complex subject matter domain. *Advances in instructional psychology*, *4* (55-175).

Schnotz, W. (2001). Conceptual Change. In D. H. Rost (Ed.), Handwo¨rterbuch Pa¨dagogische Psychologie [dictionary of educational psychology] (pp. 75e81). Weinheim: BeltzPVU.

Schnotz, W., Vosniadou, S., & Carretero, M. (Eds.). (1999). New perspectives on conceptual change. Oxford: Elsevier.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin, 86(2), 420–428. doi:10.1037/0033-2909.86.2.420

Sierpinska, A. (2000). On Some Aspects of Students' Thinking in Linear Algebra. In J-L. Dorier (Ed.),On the Teaching of Linear Algebra, pp. 209-246. Dordrecht: Kluwer Academic Publishers.

Simon, H. A. (1978). Information processing theory of human problem solving. In Estes Handbook of Learning and Cognitive Processes (Vol 5, pp. 271-295). Hillsdale: Erlbaum.

Stein, M. K., Baxter, J. A., & Leinhardt, G. (1990). Subject-matter knowledge and elementary instruction: A case from functions and graphing. *American Educational Research Journal*, *27*(4), 639-663.

Stein, M. K., & Leinhardt, G. (1989). Interpreting graphs: An analysis of early performance and reasoning. Unpublished manuscript, University of Pittsburgh, Learning Research andDevelopment Center, PA

VanLehn, K. (1999). Rule-learning events in the acquisition of a complex skill: an evaluation of CASCADE. *The Journal of the Learning Sciences*, *8*, 71e125.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. Decision sciences, 22(2), 219-240.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. Family Medicine, 37(5), 360-363.

Yerushalmy, M. (1988). Formation of algebraic concepts using multiple representation software environments. Unpublished manuscript, University of Haifa, Israel.

VITA

NAME: Noor Elagha

EDUCATION: B.A., Psychology, University of Illinois at Chicago, 2017

TEACHING EXPERIENCE: Department of Psychology, University of Illinois at Chicago: Graduate Teaching Assistant for Introduction to Research Methods in Psychology; 2019-2020

Department of Psychology, University of Illinois at Chicago: Graduate Teaching Assistant for Introduction to Psychology; 2018-2019, 2021

HONORS: Fall 2020 Provost's Graduate Research Award

PROFESSIONAL MEMBERSHIP: Psi Chi National Psychology Honor Society, 2018

PUBLICATIONS: Gane, B.D., Israel, M., **Elagha, N**., Yan, W., Luo, F., Pellegrino, J.W., (2021): Design and validation of learning trajectory-based assessments for computational thinking in upper elementary grades, Computer Science Education, DOI:10.1080/08993408.2021.1874221

Gane, B.D., **Elagha, N**., Luo, F, Liu, R, Yan, W., Strickland, C., Franklin, D., Rich, K., Pellegrino, J.W., Israel, M. (2020). Developing computational thinking assessments from learning trajectories: design approach and preliminary validity evidence. Paper accepted to the *2020 Annual Meeting of the American Educational Research Association,* San Francisco, California.

Gane, B.D., **Elagha, N**., Luo, F, Liu, R, Yan, W., Strickland, C., Franklin, D., Rich, K., Pellegrino, J.W., Israel, M. (2020). Developing computational thinking assessments for elementary students: connecting cognition, observation, and interpretation. Poster accepted to the *2020 Annual Meeting of International Conference of the Learning Sciences,* Nashville, Tennessee.

Elagha, N. (2019). The design and validation of assessments to explore the integration of computational thinking with elementary mathematics learning. Poster presented at the *2019 Learning Sciences Graduate Student Conference,* Northwestern University.

Elagha, N., Burkett, C., Blair, A. & Goldman, S. (2018, April). Effects of math anxiety on identification of discrepancies in scientific text and graph comprehension. Poster presented at the *Annual Student Research Forum*, University of Illinois Chicago, IL.

Elagha, N., Burkett, C., Blair, A. & Goldman, S. (2018, April). Effects of math anxiety on identification of discrepancies in scientific text and graph comprehension. Poster presented at the *Annual Midwestern Psychological Association Conference*, Chicago, IL.

Elagha, N., Burkett, C., Blair, A. & Goldman, S. (2017, April). Scientific text and graph comprehension through identification of discrepancies. Poster presented at *the Annual Student Research Forum*, University of Illinois Chicago, IL.