The Mathematician Educator
2020, Vol. 1, No 1, 43-57

Types of Understanding of Statistics Demonstrated by Postsecondary Students

Collette Lemieux
Mount Royal University
Canada

Olive Chapman
University of Calgary
Canada

Studies in statistics education tend to highlight students' misconceptions and difficulties in learning statistics. But, from a constructivist perspective of learning, it is important to attend to what students know or can do. In this paper, we attend to this student-oriented perspective of learning through a study that investigated postsecondary students' learning of statistics. The focus is on the types of understandings demonstrated by postsecondary business students in a first year statistics course with a pedagogical approach based on inquiry-oriented, story-based tasks. Data sources consisted of students written responses to the story-based tasks and follow-up tasks during the course. The findings indicate six ways of representing the students' understandings of specific statistics concepts in terms of what they knew and offer a way of extending the types of understanding used to make sense of students' learning of mathematics and statistics.

Keywords: understanding • statistics • postsecondary students • story-based tasks • business majors.

Introduction

In today's data-driven society, it is important for students to develop statistical literacy to become responsible citizens. This situation places statistics education in the spotlight particularly in relation to research to understand the learning and teaching of statistics at various levels of education. However, studies have tended to focus on students', and, to a lesser extent, teachers' difficulties, misconceptions or poor performance in statistics. This focus on students is highlighted in Huck's (2009) book that documents empirically demonstrated student misconceptions about statistics. While such studies contribute useful information to the field to understand learners, the focus tends to be on a deficit perspective of the learner. This perspective is likely to limit the meaningfulness of instruction to help students to develop statistical literacy. Instead, as promoted in mathematics education, understanding students' thinking from a constructivist perspective is more likely to engage students in learner-centered instruction that builds on what students know (National Council of Teachers of Mathematics, 2000). In this paper, then, instead of focusing on what students do not know, we focus on what they know and identify different types of understanding they demonstrated in learning statistics. We report on one aspect of a larger study that investigated post-secondary students' learning of statistics through inquiry-oriented story-based tasks. Specifically, we address the research question: what types of understanding of statistics concepts are demonstrated by post-secondary business students in a first year university statistics course that utilizes a pedagogical approach consisting of story-based tasks? The findings offer a way of extending the types of understanding used to make sense of students' learning of mathematics and statistics.

Related Literature

Research in statistics education has given significant attention to students' knowledge or learning of various concepts in statistics and probability. This trend is reflected in a recent (2014) special issue of Educational Studies in Mathematics on statistical reasoning. For example, studies by Garfield, Le, Zieffler, and Ben-Zvi (2014), Meletiou-Mavrotheris and Paparistodemou (2014), and Pfannkuch, Arnold, and Wild (2014) examined students' developing understanding and use of foundational statistics concepts such as randomness, samples, and sampling variability to suggest learning trajectories built upon informal inferential reasoning that are intended to support development of higher level, formal statistical thinking. However, regardless of whether studies investigated elementary or secondary school students or postsecondary (college or university) students, a commonality of findings is that students struggle with statistical concepts and demonstrate misconceptions or conceptual difficulties of many concepts. In this literature review, we focus on examples of studies on postsecondary students' learning or knowledge of statistics that have reported such findings in terms of types or levels of understanding demonstrated by students. Such studies provide a basis to determine how the types of understanding emerging from our study relate to the field.

Some studies have focused on students' understanding at the end of a statistics course. For example, Trumpower (2015) investigated the reasoning that 56 undergraduate psychology students, enrolled in an introductory behavioral statistics course, used when performing intuitive analysis of variance (IANOVA). He found that, even after completing a university course in statistics, most of the students did not demonstrate a strong understanding of variation, but some students' understanding was more advanced than others. He rated the students' understanding as developing, weak, or strong. He noted that those with a developing understanding of variation showed a somewhat different pattern of reasoning than those with weak understanding. He explained that when students with a weak understanding of variation compared data sets to determine if there was evidence of differences between the means, they tended to look at the difference between group sample means and were misled by comparing extreme values in the data sets. Only those students with a strong understanding of variation were able to realize that it is the between-group variability relative to the within-group variability that gives evidence of the differences between the means.

Mathews and Clark (2007) also examined students' understanding at the end of a course. They studied eight first year undergraduate students who recently earned an A in an introductory statistics course to investigate their understanding of the concepts of mean, standard deviation and the central limit theorem at the end of the course. Unlike Trumpower, they did not rank students' understanding but rather aimed to characterize their conceptions. They found that the students relied heavily on formulations of these concepts (i.e., the process of finding them) and had little to no conceptual understanding of any of them. For example, all of the students could compute the mean, but they confused it conceptually with mode and proportion; believed that the standard deviation is found by determining the distance between the data values in the sample or is the distance between the mean and one data value in the sample; and had no coherent understanding of the central limit theorem. The authors concluded that even students who would be deemed to have very successfully completed a postsecondary introductory statistics course may not have a strong understanding of basic statistical concepts.

While Verkoeijen, Imbos, van de Wiel, Berger, and Schmidt (2002) also considered students' understanding at the end of a course, unlike the above studies, they focused on investigating the impact of a pedagogical intervention. They evaluated a prototypical statistical learning environment based on the mental representations students had formed with respect to three statistical concepts - confidence intervals for one mean, one-sample mean *z*- and *t*-tests, and type I and II errors in statistical inference. Participants were 107 first year health science students who took part in an introductory statistics instructional cycle focusing on the basic principles of statistical inference. Findings indicated that many of the students appeared to have low levels of conceptual understanding across the three concepts. Only 40% of the participants who took the final test had acquired overall conceptual understanding on the basic principles of statistical inference. Students were able to recall statistical terms and formulas but interpretations and background knowledge of them were hardly mentioned. In addition, the examples of their incorrectly recalled elements contained some serious misconceptions.

Some studies focused on how students' understanding changed from the beginning to the end of a statistics course. For example, Dubreil-Fremont, Chevallier-Gaté, and Zendrera (2014) examined how students understanding of mean and standard deviation changed from prior to a statistics course to after. The participants were 352 undergraduate students in social sciences. Findings indicated that the students appeared to have developed a better understanding after the course. For example, more students could provide a definition of both terms at the end of the course compared to the beginning, but the definitions focused on procedure rather than concept. As another example, most of their definitions of the mean focused on how to find the mean rather than what the mean represents as a measure of center. Further, though there was improvement of both the mean and standard deviation, more students understood the mean compared to the standard deviation.

delMas, Garfield, Ooms, and Chance (2007) also compared students' understanding at the beginning and end of a statistics course. Their project included a large scale study of postsecondary students in an introductory statistics course to measure their conceptual understanding of important statistical ideas by comparing their pretest and posttest performance. Their goal was to learn more about areas in which students demonstrated improved performance, no improvement or decreased performance from beginning to end of the course, as well as an increase in students' misconceptions about particular statistical concepts. They found only modest to no increases in the students' conceptual knowledge of statistics. For example, they explained that students were correct on little more than half of the items, on average, by the end of the course and for almost all items, there was a noticeable number of students who selected the correct response on the pretest but chose an incorrect response on the posttest. They concluded that many students did not demonstrate a good understanding of much of the content covered by the test including important design principles or important concepts related to probability, sampling variability, and inferential statistics.

One study in which all participants did not take a postsecondary statistics course is Hannigan, Gill, and Leavy (2013). The aims of this study included to investigate the conceptual understanding of statistics of prospective secondary mathematics teachers. Participants were 134 prospective mathematics teachers, the majority (86%) of whom were undergraduates (in years 1–4) in a four-year physical education degree program. All of them had chosen mathematics as their elective option. There were also 19 postgraduate students in a 1-year graduate diploma program in education (mathematics teaching). The year 1 and postgraduate students had no university level statistics while the years 2 to 4 completed the introductory statistics course in their program 4 to 30 months prior to the study. Findings indicated that the conceptual knowledge of these prospective teachers was poor in some fundamental areas in statistics such as being able to properly describe the distribution of a quantitative variable and data collection and interpret box plots. In particular, they were not able to correctly identify which graphical representation best represented all the features of a distribution, why randomization was used in context, or the interpretation of centrality and variability in box plots.

The preceding sample of studies cover postsecondary students from a variety of disciplines or programs. They also cover a variety of statistics concepts. Therefore, they provide a reasonable scope of how understanding is addressed in describing students' statistical knowledge or ability. Most of the studies used conceptual understanding without defining it theoretically and at times it seemed to be used as a term to indicate knowledge of a concept. They focused on the strength (e.g., not good, poor, weak, not strong, strong), stage (e.g., low level, little to none, developing) or comparative state (e.g., modest to no increases; decreased, some or no improvement) of students' understanding of selected statistics concepts being investigated with more emphasis on students' difficulties or misconceptions. These studies also indicated that regardless of whether students' understanding was considered before or after taking a statistics course, they still demonstrated poor conceptual knowledge. The difference between such studies and our study is that we have a constructivist goal to identify and describe types of understanding of statistics concepts students were able to develop and demonstrate that make sense and not to present a deficit perspective of their knowledge or performance.

Theoretical Perspective of Types of Understanding

In mathematics education research, Hiebert and Lefevre's (1986) framework of conceptual and procedural knowledge have been used as one way of viewing types of understanding from a theoretical perspective. Hiebert and Lefevre distinguished conceptual knowledge from procedural knowledge by identifying conceptual knowledge with relationships between pieces of knowledge and procedural knowledge with having a sequential nature. They explained,

> Conceptual knowledge is characterized most clearly as knowledge that is rich in relationships. It can be thought of as a connected web of knowledge, a network in which the linking relationships are as prominent as the discrete pieces of information. Relationships pervade the individual facts and propositions so that all pieces of information are linked to some network. (pp. 3–4)

Conceptual knowledge, then, is achieved by "the construction of relationships between pieces of information" or by the "creation of relationships between existing knowledge and new information that is just entering the system" (p. 4). In contrast, for procedural knowledge, they explained,

> One kind of procedural knowledge is a familiarity with the individual symbols of the system and with the syntactic conventions for acceptable configurations of symbols. The second kind of procedural knowledge consists of rules or procedures for solving mathematical problems. (p. 7)

This indicates that procedural knowledge includes a familiarity with the symbol representation system of mathematics and knowledge of rules for manipulating symbols. Hiebert and Lefevre suggested that the relationships present in procedural knowledge are primarily sequential, that is, “It is the clearly sequential nature of procedures that probably sets them most apart from other forms of knowledge” (p. 6). They noted that, while procedural knowledge may or may not be learned meaningfully, conceptual knowledge must be learned with meaning.

A related way of viewing types of understanding in mathematics education was proposed by Kilpatrick, Swafford, and Findell (2001). They included conceptual understanding and procedural fluency as two of the strands necessary for mathematical proficiency. They defined conceptual understanding as the “comprehension of mathematical concepts, operations, and relations” (p. 5).

> [It] refers to an integrated and functional grasp of mathematical ideas. Students with conceptual understanding know more than isolated facts and methods. They understand why a mathematical idea is important and the kinds of contexts in which is it useful. They have organized their knowledge into a coherent whole, which enables them to learn new ideas by connecting those ideas to what they already know. (p.118)

In contrast, “procedural fluency refers to knowledge of procedures, knowledge of when and how to use them appropriately, and skill in performing them flexibly, accurately, and efficiently” (p. 121). Thus, while conceptual understanding focuses on comprehending connections and similarities between interrelated ideas and professional fluency includes efficiency and accuracy in basic computations, both are important to mathematical proficiency.

Based on our literature review, the preceding perspectives of types of knowledge are implied to various extent in studies on students’ knowledge or understanding of statistics. In particular, the terms conceptual knowledge and conceptual understanding have been used to highlight students’ lack of appropriate understanding or deficiencies in knowledge of statistics concepts. However, while Skemp’s (1976) perspectives of relational understanding and instrumental understanding, which are related to conceptual and procedural knowledge, have received significant attention in framing research in mathematics education, they tend to not be specifically adopted in statistics education research. Skemp described instrumental understanding as “rules without reasons” (p. 20), for example, knowing how to perform a procedure or apply a rule to solve a problem, but without knowing the reason for it or understanding how it works. Thus instrumental understanding is similar to procedural knowledge (Hiebert & Lefevre, 1986) learned without meaning. Relational understanding, on the other hand, is “knowing both what to do and why” (p. 20), for example, knowing and applying a rule while also being able to know why a rule works and connect one rule with another or being able to deduce specific rules and procedures from more general mathematical relations. This means that relational understanding occurs when procedural knowledge is linked with conceptual knowledge (Hiebert & Lefevre).

Considering our focus on *understanding* and relations through statistical *stories,* we chose to use Skemp’s (1976) language of relational and instrumental understanding, while recognizing that they are related to conceptual and procedural knowledge (Hiebert & Lefevre, 1986). Skemp’s perspectives of understanding provided a meaningful starting point for our study to investigate participants’ understanding given our use of story-based tasks with potential to engage students in instrumental and relational understanding within the same task. The story-based tasks also align with Skemp’s explanation of relational understanding, which suggests an element of relationship to the real world or contextual situations; for example, it includes being able to apply theoretical knowledge to real-world situations or new/different contextual situations and working out theory from contextual situations. In the context of statistics, students with instrumental understanding would know how to do a statistical calculation or follow a statistical procedure, but they would have difficulty interpreting the results. Further, they may know the definition of a statistics term, but could not apply the term in practice or know why the definition was appropriate. Students with relational understanding would know what statistical procedure to use and why it is appropriate to use in multiple contexts; how to perform the procedure in multiple contexts; how the procedure works; how to interpret the results (including knowing what an appropriate result would look like); the limitations of the procedure; and how the various procedures are interconnected. They would also know the reasoning behind the concepts. These ideas framed how different types of understanding students demonstrated in this study were determined.

Methodology

This study is part of a larger qualitative research project that investigated the use of *story-based tasks* to support students' learning of statistics. These tasks focus on four topics in an introductory business statistics course: descriptive statistics, informal inferential statistics, sampling distributions, and confidence intervals. The tasks, one for each topic, are written as short stories and are around 10-12 pages long. Table 1 provides a summary of one of the stories.

Table 1.
Summary of Bob's Bikes story

Bob's Bikes Story - Sampling techniques and descriptive statistics
The story of *Bob's Bikes* is that of three accounting articling students, Jolene, Franca and Bart, who are asked to determine whether an inventory system should be repaired immediately or if the cost could be deferred to a later date. The first part of the story has the characters describing sampling techniques, which they then implement. Once they have their sample, students found various descriptive statistics that they interpreted and then considered as a whole to determine whether the system should be repaired or not. In responding to the story, the students were required to explore sampling methods, find and interpret descriptive statistics, and come up with a conclusion based on the sample. *Bob's Bikes* is a comprehensive story. Jolene and Franca are the expert characters, while Bart is the novice regarding the statistics knowledge they need to make the decision about the inventory system.

Each story has a unique business-related context through which the students explored the topic. The stories are fictional, but set in realistic situations. Two type of stories were used. The first type, a *comprehensive story* (e.g., Bob's Bikes story), was used for statistical topics that were comprised of multiple statistical concepts. For example, the topic of sampling techniques and descriptive statistics involves multiple concepts, which includes various statistical measures such as mean, standard deviation and box plots. Thus, comprehensive stories were used at the end of the unit on the topic to provide the students the opportunity to consider how the separate concepts could be used together to address one problem. The second type, an *introductory story*, was used for statistical topics that covered one significant and complex statistical concept. For example, the topic of sampling distributions of sample means is abstract and involves multiple pieces that build on each other. This resulted in introductory stories being used at the beginning of the topic to motivate the learning of the concepts and to provide a concrete reason for the learning of the abstract topic. Each story had a problem that would be resolved through some form of statistical analysis. The stories were left intentionally incomplete. Within the story, the students were prompted to write dialogues between the characters which demonstrated the students' understanding of the statistics concepts involved. The students could choose to write their dialogues as one would normally see a dialogue in a story or they could write it like a script. The prompts built into the stories required that the students produce and interpret statistical measures, draw conclusions from the statistical analysis, explain their reasoning for why they chose specific statistical measures, and explain aspects of the statistical concepts covered in the stories. The students, therefore, did not simply passively read the stories but were invited to actively engage with the story. After students read each story and wrote their dialogues, they completed a follow-up task that had them apply their statistical understanding in a new and different context.

These story-based tasks were an integral part of one section of a multi-section algebra-based business statistics course intended for first-year business students at a university in southern Alberta, Canada. The course ran for four hours per week during a thirteen-week semester, with two classes per week. The course instructor was a tenured faculty member with over fifteen years of experience teaching business statistics. He worked with the first author (researcher) in piloting an initial version of the story-based tasks, which provided him with adequate experience teaching with such tasks for the study. The story-based tasks were designed only for the first three of the four units of statistics topics covered in the course.

Participants

The participants were 19 of the students enrolled in the statistics course who, at the beginning of the course, consented to participate. The majority of students were in the first year of their business programs with a few being in their second and third year. This is the only mathematics course required for the business degree and the only pre-requisite for it is grade 12 mathematics. Therefore, the majority of students did not have prior experience with the majority of the content covered in the course.

Data Collection

Data sources for this part of the project consisted of the participants' written responses to the four story-based tasks and four follow-up tasks, which were collected shortly after the due date for each in the course. They worked individually and in groups on these tasks. For the story-based tasks, their responses were in the form of dialogues embedded within the stories that demonstrated their understanding of the statistics concepts associated with the stories. The follow-up task for each story-based task was assigned to students immediately after they completed the story-based task. These tasks required that the students to apply their understanding of the statistical topics covered in the story-based task in a new and different context, which provided another way of demonstrating their understanding. These tasks are also too long to include here. One example involves the challenges of a CEO of a company that manufactures electric vehicles who is trying to break into the international market.

Data Analysis

Analysis to determine the participants' understanding of the four statistics topics (descriptive statistics, informal inferential statistics, sampling distributions, and confidence intervals) occurred in three stages. The first stage involved open-coding of the data for common ways participants demonstrated understanding (if any) for a concept; for example, the different ways that participants explained how the sampling distribution of sample means is generated and how it differs from its parent population. An example of a common way was one group's explanation of the process of resampling to create a sampling distribution in detail and making a clear distinction between the type of data in each distribution. This stage of analysis provided an initial idea of the participants' understanding.

The second stage of analysis involved coding the common ways that the participants responded to the story-based tasks based on our disciplinary knowledge and Skemp's (1976) framework for relational and instrumental understanding. Table 2 provides samples of the data from two groups of participants (groups 6 and group 4) for the story on the topic of sampling distributions that were coded as representing relational understanding or minimal instrumental understanding of how to construct a sampling distribution from a parent population.

Table 2.
Sample of relational and instructional understanding data

Sample of relational understanding – (group 6)	Sample of instrumental understanding – (group 4)
Jed: So a sampling distribution is used when the entire population is unknown, so we will take our sample of 320 and randomly select a sample of 30 from this larger sample and measure the mean. We will put these 30 back into the large sample and we will then randomly take another sample of 30 from the 320, measure the mean, and continue on until we have enough means to create a sufficient enough sampling distribution from all of our sample means. You said this process is called bootstrapping. *Reema*: Right on, it is important to note that there is a known difference between the data in the sample and the data on the sampling distribution. The sample of the 320 scooters collected were randomly sampled every 15 minutes, and then tested for peak speed which was then recorded and plotted on the curve. This is the actual sample containing raw data. The data collected from the process of bootstrapping, is the same data, except when we look at the sampling distribution, this is strictly made up from the sample means derived from bootstrapping. So it is the means of the 30 empirically sampled scooters.	*Reema*: To make sure that we get all possible samples we will use the empirical sampling distribution method. In this method we take a parent sample of 320 scooters and then take a sample from that parent sample and we will re-sample until we have sampled all of the sample. For example our sample size will be 30 and the only variable that will change will be the speed of the scooter.

Group 6's response demonstrated a thorough understanding of the process of generating a sampling distribution from the parent sample and clearly described the differences between the two types of distributions. Although their answer contained some errors (e.g., they incorrectly suggest that all sampling distributions are empirical), their overall response indicated that they had a strong understanding of the key differences and connections between the two types of distributions and, thus, relational understanding for this concept. On the other hand, group 4's response was categorized as demonstrating minimal instrumental understanding because they attempted to explain the process of resampling by referring to empirical sampling distributions but their response was vague. For example, although they stated that a sample is taken from the parent sample, what is done with each sample is unclear. Thus, even though some understanding is demonstrated about the basics of sampling distributions, there is not enough detail to suggest that the group has achieved relational understanding of sampling distributions.

Finally, the third stage of analysis involved looking across the story-based tasks for common types of understanding demonstrated. The focus was on categorizing the information from the codes found in stages 1 and 2 by looking for themes and patterns related to different levels of understanding. For example, one way that participants demonstrated instrumental understanding in all four story-based tasks was by producing statistical measures. Therefore, all of the codes that related to participants correctly producing statistical measures were grouped into the theme of algorithmic understanding. From this, six themes emerged of ways that participants demonstrated their understanding: *algorithmic understanding, terminology understanding, contextual understanding, choice understanding, basis understanding, undeveloped/developing understanding.* These ways of understanding form the basis of our findings of how the participants made sense of the four statistics topics.

Findings: Types of Understanding of Statistics Concepts

The six themes of understanding that emerged in the data analysis represent five ways in which the participants demonstrated the understandings they had developed in learning the statistics concepts and one way to account for undeveloped or developing understanding. We consider these themes to be six types of understanding: 1) algorithmic, 2) terminology, 3) contextual, 4) choice, 5) basis, and 6) undeveloped/developing understanding.

- *Algorithmic understanding* indicates that the participant could correctly follow a procedure.
- *Terminology understanding* indicates that the participant could correctly provide a definition of a statistical term in their own words.
- *Contextual understanding* indicates that the participant could correctly interpret the statistical measure in the context of the story.
- *Choice understanding* indicates that the participant could choose between at least two different statistical measures and could correctly justify that choice.
- *Basis understanding* indicates that the participant could correctly explain the reasoning behind a statistical concept.
- *Undeveloped/developing understanding* indicates that the participant did not demonstrate any of the above types of understanding and that their answer was incorrect.

We elaborate on each of these with examples to illustrate their meaning from the perspective of what students were able to do.

Algorithmic understanding involves being able to correctly follow and carry out an algorithmic procedure. This understanding occurred in situations that included students finding statistical measures (e.g., the mean, confidence intervals) by following a given/known procedure or following a specific algorithm to arrive at a correct decision or solution. For example, in comparing the *p*-value to the level of significance to correctly decide whether to reject or not reject the null hypothesis students, students were able to follow a clearly outlined algorithmic process and execute the steps to arrive at a decision. Similarly, for group tasks, multiple groups of participants were able to use predetermined process/steps to correctly produce the numerical and visual descriptive statistics of their data sets in the story-based task on descriptive statistics to determine appropriate statistical measures or make the appropriate decisions.

Terminology understanding involves being able to correctly state definitions of statistical terms or terminology. This understanding occurred in situations where the definition students provided was presented in a way that was correct but different from the way it was presented in the course resources (e.g., textbooks); that is, simply reproducing the definition from the textbook was not sufficient to demonstrate this type of understanding. For example, for a group task, one group of students (group 6) provided this definition of outliers in the story that covered the topic of descriptive statistics: "An outlier is data that lies an abnormal distance from other values." This definition demonstrates terminology understanding because, in addition to being correct, it is written in the participants' own words which differs from the wording of the definition in the textbook (i.e., "An **outlier** is an observation of data that does not fit the rest of the data." (Holmes, Illowsky, & Dean, 2016, Section 2.2, para. 4)). On the other hand, another group of students (group 4) defined an outlier as "data points that are far from the average." Although this is worded differently from the textbook, it does not demonstrate terminology understanding because it is an inadequate definition since an outlier is different from the rest of the data values and not just the average.

Contextual understanding involves being able to appropriately synthesize the contextual and the statistical information associated with a task to arrive at a correct solution of the task. It is knowing within an actual or imagined real-world contextual situations. This understanding occurred in situations that included students correctly interpreting a statistical measure in the context of the problem and appropriately using the context and statistical measures to arrive at a decision; for example, when students correctly interpreted what a confidence interval means within the context of the problem or correctly used the interpretation of a confidence interval to decide on whether a new business model should be used or not. A specific example involves Participant 4 who interpreted the *p*-value in the context of the story-based task on the topic of informal inferential statistics as follows: "Well if Aries [the dolphin] is guessing, then yes it just has a 2% chance of getting 15 or more right out of the 20." Here the participant correctly interprets the statistical measure (*p*-value) in the context of the problem by accurately stating the assumption in the conditional probability (the dolphin is guessing) and that the probability is of the evidence and even better evidence against the assumption (at least 15). On the other hand, Participant 3 interpreted the *p*-value as follows: "So Aries would have a 0.02% chance of get it right 18/20 times if he was just guessing." Participant 3 has correctly stated the condition, but has not correctly stated the even better evidence portion. These examples indicate that while Participant 4 demonstrated complete contextual understanding of the *p*-value, Participant 3 did not.

Choice understanding involves being able to choose appropriate statistical measures or models for a situation and justifying the choices. It applies to situations in which making a choice is necessary. This understanding occurred in instances, where, for example, students chose the best measure of centre for a specific set of data and justified this choice by referring to the presence or absence of outliers or to the problem context, or they chose which model best fits a situation (e.g., z or t means test) and correctly justified that choice. A specific example involves participants' (group 14) explanation of why they chose their confidence level within the story-based task that covered the topic of confidence intervals:

> The 95% confidence interval was chosen so that we could be reasonably sure of the risk Leor might be going into (investing his time and money into a business that could be

> unprofitable), while taking into consideration the risk he could be taking by not pursuing this endeavour (moving on from this business idea). Investment of time and money is more of a risk because he has a lot to lose but moving on from the idea would make him unable to reap the rewards of his work. Because of this, a 95% confidence interval seems to be the reasonable choice in this situation.

The course instructor had advised students to choose their confidence level by considering which error was more severe based on the context. Here the group correctly justified their confidence level by using the story-context to examine the competing risks (or consequences of errors) to determine that a more balanced approach of a confidence level of 95% was appropriate. On the other hand, another group of students (group 16) provided this answer: "We chose a 95% confidence interval because there was only one low outlier and this shows that they won't be risking that much money." This answer does not demonstrate choice understanding as the level of confidence should be decided prior to finding the data to avoid bias and the outlier in question was not about money.

Basis understanding involves being able to correctly explain a statistics concept in terms of both how and why. It is knowing the *basis* for the concept, where *basis* means the underlying reasons, support or foundation for the concept, or more generally, an idea, argument, or process. This understanding occurred when participants, for example, explained both how and why the sample size impacts sampling variability. A specific example is based on the following excerpt of a dialogue for the story-based task about dolphin communication that covered the topic of informal inferential statistics:

> *Emily*: The "extreme as" phrase on the applet is essentially testing the probability reaching a certain number. For example, if I were to try and find the probability of my winnings from a slots machine reaching \$50 I would also consider winning \$60 as reaching \$50. If you won \$80 and I asked you if you won \$75 what would you say?
>
> *Sam*: I would say no, I won \$80! But I kind of understand what you're saying. I won at least \$75.
>
> *Emily*: In our case, this is what we want to look at as well. If Aries [the dolphin] were to get it right 19 times out of 20, then we would count that as a success since he got it right at least 17 times.

In relation to this dialogue, Participant 17 correctly explained "why the 'at least 17/20' probability is more meaningful in helping us answer our research question" instead of "exactly 17". That is, the participant explained the reasoning behind the "or even better evidence against the null hypothesis" portion of the definition of the *p*-value within the story context. Participant 17 correctly explained that higher scores were even better evidence of communication between dolphins by relating the concept to the analogy of gambling, then returned to the story-context and explained how the analogy related to the statistical concept. The dialogue, therefore, demonstrated basis understanding as the participant correctly explained the underlying reasons (basis) for the concept. On the other hand, Participant 16 did not demonstrate basis understanding of this concept in explaining: "We test the probability of getting at least 17 out of 20 because as the value increases the probability of meeting these restrictions will only decrease." This explanation describes what happens rather than why it happens and lacks clarity regarding what "value" represents since if "as the value increases" means that as more values are added to the probability (i.e., 17, 18, 19, 20), then it is incorrect to say the probability decreases.

Undeveloped/developing understanding involves being able to partially or incorrectly demonstrate a statistics concept or process. It occurred in situations where students' thinking did not demonstrate any of the above five types of understanding or students offered vague, unclear, or inadequate knowledge of the concept or procedure. Some examples of this 'understanding' are the counter examples given above for terminology, contextual, choice, and basis understanding. The counter terminology understanding example involved group 4's definition of an outlier as "data points that are far from the average," which is inadequate because as an outlier is different from the rest of the data values and not just the average. The counter contextual understanding example involved Participant 3 whose interpretation of the *p*-value in the context of the story-based task on the topic of informal inferential statistics was partially correct. The counter choice understanding example involved group 6 who provided the answer: "We chose a 95% confidence interval because there was only one low outlier and this shows that they won't be risking that much money." This answer is inadequate since the level of confidence should be decided prior to finding the data to avoid bias and the outlier in question was not about money. The counter basis understanding example involved participant 16's response that suggested if "as the value increases" means that as more values are added to the probability (i.e., 17, 18, 19, 20), then it is incorrect to say the probability decreases. Another example (connected to the basis understanding example situation above) involves participant 2 who wrote:

> Exactly means that only value which in our case is 17/20 however… at least means that is a combined set of value with all the values above it as well. All the 18/20, 19/20 & 20/20 probabilities have been added.

This explanation has the correct meaning of "at least" (i.e., what it is), but does not explain why it is "at least" as required by the task. All of these examples suggest that the students had not developed or were in the process of developing understanding of the statistics concepts involved.

Discussion

This study focused on identifying the types of understanding of specific statistics concepts that students demonstrated through the use of story-based tasks. Six types were identified: algorithmic, terminology, contextual, choice, basis, and undeveloped/developing understanding. The first five represent the nature of the knowledge the students had developed and demonstrated in their work, that is, these were their constructed or *developed* understanding. The sixth type represents *undeveloped/developing* understanding that occurred when students were only able to provide vague, unclear, or inadequate knowledge of the concept or procedure, that is, from a constructivist perspective, students were at a stage where they had not fully developed or were in the process of developing their understanding of a concept or way of expressing it with clarity.

The story-based tasks provided students with the opportunity to develop the different types of understanding, but students varied in which type they developed. Though all six types of understanding emerged from the participants' work, some of them were more prevalent than others. For example, most (over 75%) participants demonstrated algorithmic, terminology and contextual understanding for most concepts in the four selected topics (descriptive statistics, informal inferential statistics, sampling distributions, and confidence intervals). Choice understanding was demonstrated by most participants for the topics of descriptive statistics and informal inferential statistics. Basis understanding was only demonstrated by most participants for the topic of sampling distributions of sample means. Additionally, most participants who demonstrated choice, contextual or basis understanding in one task or context continued to demonstrate that type of understanding in a new task or context. A more detailed description of these findings are beyond the scope of this paper. But these findings suggest that students could hold their knowledge in terms of a complex combination of these types of understandings that could get masked or simplified when they are considered as instrumental/relational or procedural/conceptual combination. For example, the expanded types made it possible to recognize that a student may know how to use a formula to produce a statistical measure, but cannot provide a definition of statistical measure, thus demonstrating one aspect of instrumental understanding, but not all aspects. Similarly, a student may successfully interpret the results of a statistical investigation but may not be able to explain how the statistical procedure works, thus, demonstrating one aspect of relational understanding but not the other. Together, then, the six types of knowledge provided a more appropriate, multi-faceted, and positive basis to represent the nature of students' understanding of statistics.

While we started with Skemp's (1976) instrumental and relational understandings, we found that expanding them into six specific types of developed and undeveloped/developing knowledge provided a more meaningful basis to represent and highlight what students know. However, the five types of developed understanding can be viewed as a decomposition of Skemps' definitions of instrumental and relational understanding as well as Hiebert and Lefevre's (1986) definition of procedural and conceptual knowledge and Kilpatrick et al.'s (2001) definition of procedural and conceptual understanding in terms of knowing *how*, *what*, *meaning/use*, *which*, and *why/reason*. For the five types of understanding, these can be viewed as *how* for algorithmic and *what* for terminology, *meaning/use* for contextual, *which* for choice, and *why/reason* for basis. For example, 'how' indicates knowing steps, 'what' indicates knowing definition, 'meaning/use' indicates ability to interpret and use concept, 'which' indicates ability to choose and justify, and 'why/reason' indicates knowing reason underlying the concept. Thus, algorithmic and terminology are related to instrumental understanding (e.g., knowing how to do an algorithm and what a term means) while choice, contextual, and basis are related to relational understanding (e.g., knowing which statistical measure or model to use and *why* it is appropriate to use; knowing *why* the mean is a measure of central tendency). For algorithmic understanding, there was no indication that participants understood why the steps of the algorithm were appropriate. For terminology understanding, while the participants were able to describe the definition in their own words, this does not necessarily indicate that they understood why it was appropriate. For example, just because they can write the definition of the mean in their own words does not imply that they understand why the mean is a measure of central tendency. In contrast, for choice, contextual, and basis understanding, participants demonstrated knowledge of 'why'. For example, choice included justifying why a concept/measure/model was appropriate to use, contextual included demonstrating what the result means and why it is important or relevant to understanding the problem (i.e., why and how a statistical measure can provide insight into the context and problem being investigated), and basis included explaining the reasoning behind a concept.

As noted in our literature review, studies tend to approach students' learning of statistics from a deficit perspective that highlights their struggles, misconceptions, or weak conceptions of statistics concepts. In contrast, we defined the six types of knowledge in a way that highlights what students know or can do. This is consistent with viewing learning from a constructivist perspective and the importance for pedagogical interventions that provides opportunities to build on students' developed or developing understanding during a statistics course. They therefore have implications for supporting research and teaching.

Implications

The findings suggest a model of understanding that could be useful in exploring students' learning, not only of statistics, but also mathematics in general. For example, in calculus, students can demonstrate algorithmic understanding of derivatives by following steps to find a derivative or choice understanding by choosing between different integration techniques and justifying why the choice would be appropriate. Table 3 provides a general interpretation of components of this model.

Table 3.
Types of understanding

Types of understanding	Meaning
Algorithmic	Able to correctly perform a procedure.
Terminology	Able to correctly provide a definition of a concept in one's own words.
Contextual	Able to correctly interpret a concept in a context.
Choice	Able to correctly choose appropriate strategy/technique/measure *and* explain why it is appropriate.
Basis	Able to correctly explain the reasoning behind a concept in depth.
Undeveloped/developing	Able to provide vague, unclear, or inadequate knowledge of a concept or procedure

This model is useful to researchers in statistics education to meaningfully frame studies on students' learning since, as our literature review suggests, most of the studies used conceptual understanding without defining it theoretically and at times it seemed to be used as a term to indicate knowledge of a concept. Researchers of both statistics and mathematics education could further investigate this model and its application to make sense of students' learning of mathematics. Teacher educators could use it to help prospective teachers to interpret students' thinking or knowledge and how to engage students meaningfully in their learning. Practicing teachers could use it to engage in formative assessment and to prompt students in specific ways by acknowledging what they know and building on it to help them to further develop undeveloped/developing understanding. To conclude, this model is not intended to be an absolute way of interpreting understanding but an emerging idea that could be further researched and developed to produce a practical, learner-centered, humanistic way of representing different stages of students' learning of statistics or mathematics concepts.

References

delMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, *6*(2), 28–58. https://www.stat.auckland.ac.nz/~iase/serj/ SERJ6(2)_delMas.pdf

Dubreil-Frémont, V., Chevallier-Gaté, C., & Zendrera, N. (2014). Students' conceptions of average and standard deviation. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the ninth international conference on teaching statistics*. International Statistical Institute. http://icots.info/9/proceedings/pdfs/ICOTS9_C142_DUBREILFREMONT.pdf

Garfield, J., Le, L., Zieffler, A., & Ben-Zvi, D. (2014). Developing students' reasoning about samples and sampling variability as a path to expert statistical thinking. *Educational Studies in Mathematics, 88*, 327-342. https://doi.org/10.1007/s10649-014-9541-7

Hannigan, A., Gill, O., & Leavy, A. M. (2013). An investigation of prospective secondary mathematics teachers' conceptual knowledge of and attitudes towards statistics. *Journal of Mathematics Teacher Education, 16*, 427–449. https://doi.org/10.1007/s10857-013-9246-3

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: A case of mathematics* (pp. 1-26). Lawrence Erlbaum Associates.

Holmes, A., Illowsky, B., & Dean, S. (2016, August 28). *Introductory business statistics*. OpenStax. https://legacy.cnx.org/content/m53808/latest/?collection =col11969/1.5

Huck, S. H. (2009). *Statistical misconceptions*. Psychology Press.

Kilpatrick, J., Swafford, J. & Findell, B. (Eds.) (2001). *Adding it up: Helping children learn mathematics*. National Academy Press.

Mathews, D. & Clark, J. M. (2007). *Successful students' conceptions of mean, standard deviation, and the central limit theorem*. http://www1.hollins.edu/faculty/clarkjm/stats1.pdf

Meletiou-Mavrotheris, M., & Paparistodemou, E. (2014). Developing students' reasoning about samples and sampling in the context of informal inferences. *Educational Studies in Mathematics, 88*, 385-404. https://doi.org/10.1007/s10649014-9551-5

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. National Council of Teachers of Mathematics.

Pfannkuch, M., Arnold, P., & Wild, C. J. (2014). What I see is not quite the way it really is: Students' emergent reasoning about sampling variability. *Educational Studies in Mathematics, 88*, 343-360. https://doi.org/10.1007/s10649-014-9539-1

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching, 77*, 20-26.

Trumpower, D. L. (2015). Aspects of first year statistics students' reasoning when performing intuitive analysis of variance: Effects of within- and between-group variability. *Educational Studies in Mathematics, 88*, 115–136. https://doi.org/10.1007/s10649-014-9574-y

Verkoeijen, P. P. J. L., Imbos, Tj., van de Wiel, M. W. J., Berger, M. P. F., & Schmidt, H. G. (2002). Assessing knowledge structures in a constructive statistical learning environment. *Journal of Statistics Education, 10*(2). https://jse.amstat.org/v10n2/verkoeijen.html

---

Authors

Collette Lemieux (Corresponding author), Mount Royal University, 4825 Mount Royal Gate SW, Calgary, Alberta, Canada T3E 6K6. clemieux@mtroyal.ca

Olive Chapman, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N1N4. chapman@ucalgary.ca