British Journal of Education
Vol.10., Issue 8, pp. 1-17, 2022
Print ISSN: 2054-6351(Print)
Online ISSN: 2054-636X (Online)

Enhancing Undergraduate Mathematics Students' Conceptual Knowledge of the Confidence Interval for the Population Mean

[1]Charles Kojo Assuah, [2]Thomas Mensah–Wonkyi, [3]Matilda Sarpong Adusei, [4]Grace Abedu, & [5]Stephen Ghunney
[1,2,3]Department of Mathematics Education, University of Education, Winneba, P.O. Box 25, Winneba-Ghana
[4]Unipra South Inclusive Basic Schools, Winneba Basic Schools, Winneba
[5]A.M.E Zion Girls Senior High School, Winneba Senior High Schools, Winneba

Citation: Charles Kojo Assuah, Thomas Mensah–Wonkyi, Matilda Sarpong Adusei, Grace Abedu, & Stephen Ghunney (2022) Enhancing Undergraduate Mathematics Students' Conceptual Knowledge of the Confidence Interval for the Population Mean, *British Journal of Education*, Vol.10., Issue 8, pp. 1-17

ABSTRACT: *This study illustrates the teaching strategies that undergraduate mathematics lecturers might employ to improve their students' conceptual knowledge of the confidence interval for the population mean. The study employed an action research method. It allowed the researcher to deepen his knowledge of the subject matter by planning, acting, evaluating, refining, and learning from this experience (Koshy, 2010). The participants consisted of sixty (60) level 200 mathematics students, who were randomly selected from a cohort of mathematics students from a mid-sized university in Ghana. The students completed the tasks assigned to them in their various groups by working collaboratively together, with their lecturer helping the process. They determined the particular theorem to apply in any given situation and applied the confidence interval formula to calculate the confidence intervals. The results indicated that collaborative learning, combined with effective instructional methods, improves students' conceptual knowledge of the confidence interval. An implication of this study is that students' prior experiences act as a catalyst to enhance their conceptual knowledge. Thus, students who have a conceptual grasp of sampling techniques can conceptualise confidence intervals with ease. The study concludes that students should thoroughly understand the definitions and theorems relating to a statistical concept before they examine concrete examples relating to confidence intervals.*

KEYWORDS: Conceptual knowledge, confidence interval, action research method, the population mean, instructional methods.

INTRODUCTION

Learning statistics relies on existing knowledge structures and improves with experience and teaching. These knowledge structures have a significant impact on students' conceptual knowledge

@ECRTD-UK: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

(see, e.g., Smith, diSessa, & Roschelle, 1994). It's vital to figure out whether students have a proper statistical conceptual understanding or whether they have any misconceptions. Understanding statistical ideas are crucial since it allows you to make better statistical decisions (e.g., Garfield, 2003). To achieve this, a transition from rote memorization, which requires computation, to conceptual learning is required (e.g., Moore, 1997). The understanding of the concepts and relationships that underpin a domain is referred to as conceptual knowledge (Hiebert & Lefevre, 1986). This knowledge allows students to think flexibly (e.g., Jones, Jones, & Vermette, 2011), and transfer knowledge to novel problems (e.g., Bude, Imbos, van de Wiel, & Berger, 2011; Bude, van de Wiel, Imbos, & Berger, 2010; Paas, 1992), choose which type of analysis to use (e.g., Bude et al., 2010; Bude et al (e.g., Garfield & Chance, 2000). Regrettably, many statistics classes do not emphasise a high level of conceptual understanding (delMas, Garfield, Ooms, & Chance, 2007; Meletiou-Mavrotheris & Lee, 2002; Pfannkuch, Wild, & Parsonage, 2012). As a result, learning concepts is more difficult than learning methods (Leppink, Broers, Imbos, van der Vleuten, & Berger, 2012).

LITERATURE REVIEW

Conceptual knowledge of *CIs* necessitates the ability of an individual to comprehend the estimated intervals and relate them to other statistical concepts. Although many statistical concepts are related to CIs in various ways, complete knowledge of *CIs* appears to require a few basic concepts. Lockwood, Yeo, Crooks, Nathan, and Alibali outline key features that tend to be particularly important when thinking about CIs (2014). These features include (a) understanding the relationship between sample and population mean, (b) understanding the idea of *CI* (i.e., 90% vs. 95% *CI*), (c) understanding how various factors (e.g., sample size, sample variability) affect the width of CI, (d) understanding what can be learnt about future replications from *CIs*, and (e) understanding how to evaluate *CIs* effectively. Unfortunately, previous research indicates that some researchers and students have difficulty grasping *CI* concepts (e.g., Coulson, Healey, Fidler, & Cumming, 2010; Cumming, 2006; Henriques, 2016).

A number of misconceptions have been identified in previous studies (Castro Sotos, Vanhoof, Van den Noortgate, & Onghena, 2007; Cumming & Maillardet, 2006; Fidler, 2006; Grant & Nathan, 2008; Greenland et al., 2016; Henriques, 2016). A few of these misconceptions are: “There is a 95% chance that the true population mean falls within the confidence interval.” (FALSE), and “the mean will fall within the confidence interval 95% of the time.” (FALSE). In the past, researchers (e.g., Cumming, Williams, & Fidler, 2004; Hoekstra, Rouder, Morey, & Wagenmakers, 2014), graduate students (e.g., Grant & Nathan, 2008; Hoekstra et al., 2014), and undergraduate students (e.g., Grant & Nathan, 2008; Hoekstra et al., 2014) have all expressed concerns about *CIs* (e.g., Fidler, 2006; Henriques, 2016; Reaburn, 2014).

CIs are used for both estimation and hypothesis testing, and they necessitate a firm grasp of the concept. In fact, *CIs* are particularly important because of their ability to make statistical inferences (e.g., American Psychological Association, 2010). In *CI* interpretation, the distinction between samples and populations is crucial. *CIs* are a kind of inferential statistic that allows for the generalisation of sample data to the population. These types of generalisations necessitate a wider understanding of samples. Students' understanding of sample means concepts appear to be a substantial impediment to their understanding of *CIs* (Fidler, 2006).

This study was underpinned by Lev Vygotsky's social constructivism learning theory. According to this view, language and culture are the frameworks through which humans experience, communicate, and comprehend reality. Vygotsky claims that learning concepts are transmitted through language and interpreted and comprehended by experience and interactions within a cultural context. He believes that social connection is critical for long-term development and that social learning contributes to cognitive growth. Social constructivism emphasises the collaborative nature of learning when led by a facilitator or in collaboration with other students.

This method supports the establishment of opportunities for students to collaborate with their lecturers and peers in the construction of knowledge. It views the social aspect of learning, as well as the use of dialogue, collaboration, and application of information, as a significant aspect of learning and a means of accomplishing learning objectives. In an educational or training institution, this could be accomplished through group discussion, teamwork, or any other form of instructional contact. As they interact with others, students get the knowledge and experience they need to solve problems. Through social constructivism, students can learn collaboratively by forming various groupings and interactive ways, such as whole-class debates, small-group discussions, or working in pairs on specific projects or tasks.

The ultimate goal of this method is to develop independent thinkers who can enhance their knowledge and skills by utilising a range of resources both within and outside the classroom (Kuncel, 2008). Lecturers who combine a collaborative teaching method with an active study habit might be able to accomplish this. By inspiring students, this strategy would entice them to participate in a lesson. It piques students' interest and motivates them to tackle issues. If active study habits are implemented in the classroom, students would be self-reliant and competent in their class participation. The purpose of this study was to illustrate the teaching strategies that undergraduate mathematics lecturers might employ to improve their students' conceptual knowledge of the confidence interval for the population mean. The following research questions guided the study: (1) What instructional methods can lecturers use to teach students the concept of *CI*? (2) How do these methods aid students to understand *CI* more conceptually?

METHOD

Research Design

The study used an action research methodology. It is a method of inquiry in which a teacher or lecturer builds his or her understanding of a research problem by planning, acting, evaluating, refining, and learning from the experience (Koshy, 2010). It provides a way for teachers to engage in a more intentional, substantive, and ~~a~~ critical reflection that is documented and analysed in order to enhance their teaching. It collects evidence in order to implement change in practice; it is participatory and collaborative, and it is carried out by teachers who have ~~a~~the same goal. It is context-oriented and encourages participants to build a reflective practice based on their interpretations. It is iterative in nature, with plans being produced, implemented, amended, and then implemented again, allowing for continuous reflection and revision. As the action unfolds and takes place, the findings of an action research design emerge. They are not, however, decisive or absolute, but rather continuing (Koshy, 2010).

Participants and Setting

The participants consisted of sixty (60) (forty-five (45) males and fifteen (15) females) level 200 mathematics students from a mid-sized university in Ghana. The students were admitted from the public senior high schools across the country. They had all enroled and passed two (2) introductory statistics courses at level 100, with a grade of C or better. The average age of the students was twenty (20) years and two (2) months.

Table 1 Diagnostic Test for the Confidence Interval for the Population Mean

Item	Statement	Frequency		Correct Answer
		True (T)	False (F)	
1	A 95% confidence interval is the interval in which one is 95% certain that it contains the population mean.	25	35	T
2	If one repeatedly takes a sample of size *n* from a population and constructs a 95% confidence interval each time, 95% of these intervals should contain the population mean.	20	40	T
3	A confidence interval gives one the range of possible values for the sample mean.	37	23	F
4	If one were to conduct an infinite number of experiments exactly like the original experiment, a 95% confidence interval would contain 95% of the sample means for these experiments.	41	19	F
5	A confidence interval gives one the range of the individual scores	44	16	F
6	A confidence interval gives you the range of individual scores within one standard deviation of the population mean.	36	24	F
7	A 95% confidence interval indicates that there is a 95% chance that the sample mean equals the population mean.	38	22	F
8	A 95% confidence interval means the probability that the mean is within the interval is 95%.	36	24	F

Table 1 shows the results of a diagnostic test for the confidence interval for the population mean. The test consists of the meaning, interpretation, and application of the confidence interval (*CI*) for the population mean, and was given to the students to assess their initial understanding. The results indicated that for each of the eight (8) statements, more than half of the students selected a wrong answer. Among the students who chose the correct answer, they were unable to justify why they chose that answer. An interview conducted among the students revealed that their inability to answer the questions correctly stemmed from the fact that they lacked conceptual knowledge of the *CI*, thereby making it difficult to apply the concept. As they struggled to understand the *CI*, they internalized their conception, leading to a series of misconceptions. If these misconceptions are not corrected, they could affect their understanding in advanced statistics courses.

Instructions and Tasks

The study was divided into twelve (12) groups, each with five (5) students. All of the students came together as a class, and their lecturer explained their duties and responsibilities as they worked in their groups. He encouraged them to take an active role in the learning process (i.e., they were allowed to come up with their ideas, questions, and definitions). He highlighted that successful learning occurs through social contact; thus, students must solve problems and research and explore circumstances in groups in order to reach conclusions. While the lecturer worked as a guide or facilitator through a regulated and structured process of peer interaction, the students acquired or produced knowledge on their own. He guided them with excellent directed questions and stimulated discussions with demonstrations of certain topics, situations, or scenarios. The lecturer (i) created a social constructivist classroom environment to encourage group interaction (ii) discouraged competition while encouraging collaboration and sharing of experience among students (iii) valued the students' opinions or contributions (iv) provided the necessary resources and guidance to prompt the students to construct knowledge in the desired direction (v) guaranteed that students felt comfortable asking and answering questions, interacting with one another, and contributing to group discussions. (vi) ensured that both brilliant and less brilliant students learnt from one other, and (vii) gave timely support where needed. He stated that any success made by a group in the learning process was attributed to the contributions and collaboration of all members of the group rather than individual members. For the sake of this study, the groups were given the following pseudonyms: G1 = a member of group 1; G2 = a member of group 2; G3 = a member of group 3; G4 = a member of group 4; G5 = a member of group 5; G6 = a member of group 6; G7 = a member of group 7; G8 = a member of group 8; G9 = a member of group 9; G10 = a member of group 10, G11: a member of group 11; and G12 = The lecturer mediated the students' conceptual knowledge of the *CI* in the vignette below:

L: Students, the theorems below will help you conceptually grasp the *CI*.

Theorem 1: If random samples of size n are selected from a normal population with mean μ and variance σ^2, the sampling distribution of the sample mean $\bar{x}$, will be normally distributed, with $\mu_{\bar{x}} = \mu$ and variance $\sigma_{\bar{x}}{}^2 = \frac{\sigma^2}{n}$.

Theorem 2: (The Central Limit Theorem [*CLT*]): If random samples of size n are selected from a population with mean μ and variance σ^2, the sampling distribution of the sample mean $\bar{x}$ will be

@ECRTD-UK: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

approximately normally distributed, with mean $\mu_{\bar{x}} = \mu$, and variance ${\sigma_{\bar{x}}}^2 = \frac{\sigma^2}{n}$. To apply theorem 2 successfully, the following conditions should be noted:
If the sample size is large ($n \geq 30$), the sampling distribution of $\bar{x}$ is approximately normal, regardless of the shape of the population.
If the sample size is small ($n < 30$), the sampling distribution of $\bar{x}$ is approximately normal, provided that the shape of the population is not drastically different from normal.
If the sample is small ($n < 30$) and the shape of the population resembles an inverted normal curve, the CLT does not apply.

Theorem 3: If the conditions of theorems 1 and 2 are satisfied, the quantity $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ is a z score and represents the value of a standard normal random variable. When dealing with sample sizes where $n \geq 30$, the sample standard deviation s replaces σ ($\sigma \approx s$).

Theorem 4: If random samples of size n are selected from a normal population with mean μ and variance σ^2(unknown), then the quantity $t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$ is the value of a random variable having t distribution with $n-1$ degrees of freedom. In small sample situations ($n < 30$), we cannot replace σ with s because the quantity $\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$ does not produce values of the standard normal z under these conditions. Gosset (1908) called this distribution the t distribution or student's t distribution.
Case 1: If the conditions of Theorem 1 and 2 are met, a $(1-\alpha) \times 100\%$ interval for μ is given by $\bar{x} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{x} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$, where $n \geq 30$. An upper bound on the error of estimation, $\mu - \bar{x}$, is given by the quantity $E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$. The sample size needed in order to be $(1-\alpha) \times 100\%$ confident that the error of estimation does not exceed E is given by $n = \left[\frac{z_{\alpha/2}\sigma}{E}\right]^2$.
Case 2: If the conditions of Theorem 4 are met, a $(1-\alpha) \times 100\%$ confidence interval for μ is given by $\bar{x} - t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{x} + t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$, where $n < 30$ and the degrees of freedom are given by $\nu = n - 1$.
Tables 2 and 3 show the critical values of the normal distribution and critical values of the t distribution, respectively. The Margin of Error, $ME = \bar{x} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$ or $ME = \bar{x} + t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$. Table 2 and 3 indicate the critical values of the normal distribution and critical values of t distribution respectively.

Table 2 Critical values of the normal distribution

Confidence Level	Critical (z) Value Used in Confidence Interval Calculation
50%	0.67449
75%	1.15035
90%	1.64485
95%	1.95996
97%	2.17009
99%	2.57583
99.9%	3.29053

Table 3 Critical values of the t Distribution

	α	
ν	0.025	0.05
1	12.706	6.314
2	4.303	2.920
3	3.182	2.353
4	2.776	2.132
5	2.571	2.015
6	2.447	1.943
7	2.365	1.895
8	2.306	1.860
9	2.262	1.833
10	2.228	1.812
11	2.201	1.796
12	2.179	1.782
13	2.160	1.771
14	2.145	1.761
15	**2.131**	**1.753**
16	**2.120**	**1.746**

Source: Fisher (1973). *Statistical methods for research workers* (14th ed.). Hafner Press

Question 1*:* Given that the mean starting salary of health-care professionals in a country is μ. An employment agency wishes to estimate this mean starting salary of these health-care professionals. A random sample of the starting salaries of 100 health-care professionals in the country produced a sample mean of $\bar{a}$ with a standard deviation of σ.
Construct (i) 90% (ii) 95% confidence intervals for the true mean starting salary of all health-professionals in the country.
(iii) Do the intervals really contain the population parameter μ?

(iv) Practically demonstrate the true meaning of the 90% and 95% confidence intervals?

Steps to Conceptual Understanding of the CI

L: All the groups should discuss ideas, concepts and solutions together as a group, and every member's idea should be welcomed. When you have built a consensus on an idea or answer, or a solution, it would be deemed as being representative of that group.
L: Groups, determine whether either the Theorem 1 or Theorem 2 applies to this question.
G4: The shape of the population is unknown, and mean of the population is to be estimated. Standard deviation = σ. Since the shape of the population is unknown, and the sample size is $n = 100$, theorem 2 applies.
L: G9, describe the sampling distribution of $\bar{a}$.
G9: The sampling distribution of $\bar{a}$ follows from the theorem, its approximately normal, with mean, $\mu_{\bar{a}} = \mu$, and standard deviation $\sigma_{\bar{a}} = \frac{\sigma}{\sqrt{100}} = \frac{\sigma}{10}$. Therefore, the quantity $z = \frac{\bar{a}-\mu}{\sigma/\sqrt{n}}$ produces a value of a standard normal random variable.
L: All groups should construct a 90% CI by using the formula for *CI*.
G10: $1 - \alpha = 0.90 \Longrightarrow \alpha = 0.10, and\ so\ \frac{\alpha}{2} = 0.05$. From table 1, $z_{0.05} = 1.64485$. By substituting $n = 100$, $z = 1.64485$ into the interval, $\bar{a} - z_{0.05}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{a} + z_{0.05}\left(\frac{\sigma}{\sqrt{n}}\right)$, gives $\bar{a} - (1.64485)\left(\frac{\sigma}{\sqrt{100}}\right) < \mu < \bar{a} + (1.64485)\left(\frac{1{,}075}{\sqrt{100}}\right)$, which simplifies to $\bar{a} - 0.164485\sigma < \mu < \bar{a} + 0.164485\sigma$.
L: G12, what does this interval represent?
G12: This interval represents 90% confidence for the estimated population mean μ. This means that one is 90% confident that the interval from $\bar{a} - 0.164485\sigma$ to $\bar{a} + 0.164485\sigma$ contains the value of μ.
L: Does this interval contain the estimated population mean?
G6: It may contain the estimated population mean, which means it is not a certainty.
L: Can we therefore say that the probability that the interval contains the estimated population mean is 0.90?
G3: Yes
L: Do the rest of the groups agree with G3?
G9: We disagree because either the interval contains the estimated population mean, or it does not. So, the probability is either 0 or 1.
L: Great G9, you're right with that explanation.
L: Groups, what practical steps will you take to demonstrate your understanding of a 90% *CI*?
G4: We can construct such intervals many times, but 90% of them will contain the estimated population mean and 10% will not.
L: Good work done, G4. Could you explain using the question above?
G4:10 samples of the salaries of 100 health-care professionals should randomly be selected from the population. The mean of each of the 10 samples can be calculated using the formula $\bar{a} = \frac{\sum a}{n}$. The means of the samples with their corresponding confidence intervals are indicated in table 4.

@ECRTD-UK: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

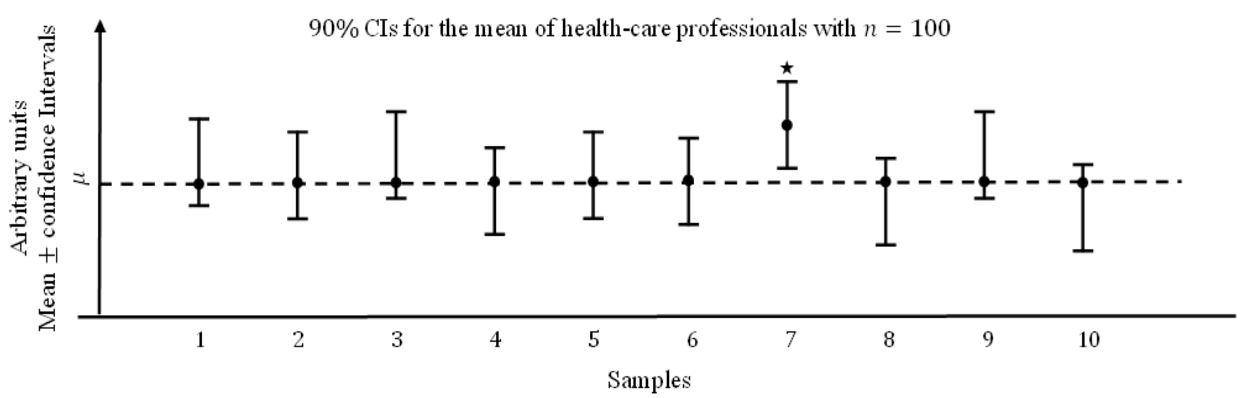
British Journal of Education
Vol.10., Issue 8, pp. 1-17, 2022
Print ISSN: 2054-6351(Print)
Online ISSN: 2054-636X (Online)

L: It is true that with 10 samples, one interval will not contain the estimated population mean, as demonstrated in table 4. Figure 1 indicates the 90% *CI* for the mean of health-care professionals with n = 100.

Table 4 Sample means with 90% *CIs*

Sample Mean	90% Confidence intervals	Is the mean within the interval?	Width of the Interval
$\bar{a}$	$\bar{a} - 0.164485\sigma,\ \bar{a} + 0.164485\sigma$	YES	0.32897σ
$\bar{b}$	$\bar{b} - 0.164485\sigma,\ \bar{b} + 0.164485\sigma$	YES	0.32897σ
$\bar{c}$	$\bar{c} - 0.164485\sigma,\ \bar{c} + 0.164485\sigma$	YES	0.32897σ
$\bar{d}$	$\bar{d} - 0.164485\sigma,\ \bar{d} + 0.164485\sigma$	YES	0.32897σ
$\bar{e}$	$\bar{e} - 0.164485\sigma,\ \bar{e} + 0.164485\sigma$	YES	0.32897σ
$\bar{f}$	$\bar{f} - 0.164485\sigma,\ \bar{f} + 0.164485\sigma$	YES	0.32897σ
$\bar{g}$	$\bar{g} - 0.164485\sigma,\ \bar{g} + 0.164485\sigma$	NO	0.32897σ
$\bar{h}$	$\bar{h} - 0.164485\sigma,\ \bar{h} + 0.164485\sigma$	YES	0.32897σ
$\bar{\imath}$	$\bar{\imath} - 0.164485\sigma,\ \bar{\imath} + 0.164485\sigma$	YES	0.32897σ
$\bar{\jmath}$	$\bar{\jmath} - 0.164485\sigma,\ \bar{\jmath} + 0.164485\sigma$	YES	0.32897σ

Figure 1 90% *CI* for the mean of health-care professionals with n = 100



L: All the groups should construct a 95% *CI* for estimated population mean.
G5: For the 95% confidence intervals, 20 samples of the salaries of 100 health-professionals were randomly selected from the population and the *CI* calculated using the same formula. An interval may be obtained as follows: $\bar{a} - (1.95996)\left(\frac{\sigma}{\sqrt{100}}\right) < \mu < \bar{a} + (1.95996)\left(\frac{\sigma}{\sqrt{100}}\right)$, which gives $\bar{a} - 0.195996\sigma < \mu < \bar{a} + 0.195996\sigma$. The means of the 20 samples, with their corresponding confidence intervals are indicated in table 5.
L: How would you practically demonstrate the 95% *CI*?

@ECRTD-UK: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

British Journal of Education
Vol.10., Issue 8, pp. 1-17, 2022
Print ISSN: 2054-6351(Print)
Online ISSN: 2054-636X (Online)

G11: When the 20 CIs are constructed, 19 of them will contain the population mean, whilst 1will not.
L: Another way of looking at the 95% *CI* is, if the intervals of many randomly selected samples are constructed, 95% of the intervals will contain the population mean, whilst 5% will not. Figure 2 indicates the 95% *CI* for the mean of health-care professionals with n = 100

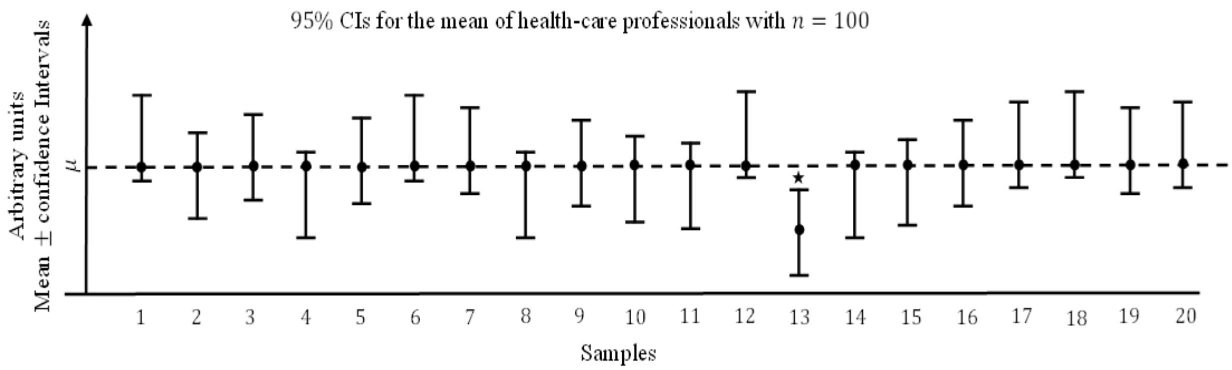
Table 5 Sample means with 95% *CIs*

Sample mean	95% Confidence Intervals	Is the mean within the interval?	Width of the Interval
$\bar{a}$	$\bar{a} - 0.195996\sigma, \bar{a} + 0.195996\sigma$	YES	0.391992 σ
$\bar{b}$	$\bar{b} - 0.195996\sigma, \bar{b} + 0.195996\sigma$	YES	0.391992 σ
$\bar{c}$	$\bar{c} - 0.195996\sigma, \bar{c} + 0.195996\sigma$	YES	0.391992 σ
$\bar{d}$	$\bar{d} - 0.195996\sigma, \bar{d} + 0.195996\sigma$	YES	0.391992 σ
$\bar{e}$	$\bar{e} - 0.195996\sigma, \bar{e} + 0.195996\sigma$	YES	0.391992 σ
$\bar{f}$	$\bar{f} - 0.195996\sigma, \bar{f} + 0.195996\sigma$	YES	0.391992 σ
$\bar{g}$	$\bar{g} - 0.195996\sigma, \bar{g} + 0.195996\sigma$	YES	0.391992 σ
$\bar{h}$	$\bar{h} - 0.195996\sigma, \bar{h} + 0.195996\sigma$	YES	0.391992 σ
$\bar{\imath}$	$\bar{\imath} - 0.195996\sigma, \bar{\imath} + 0.195996\sigma$	YES	0.391992 σ
$\bar{\jmath}$	$\bar{\jmath} - 0.195996\sigma, \bar{\jmath} + 0.195996\sigma$	YES	0.391992 σ
$\bar{k}$	$\bar{a} - 0.195996\sigma, \bar{k} + 0.195996\sigma$	YES	0.391992 σ
$\bar{l}$	$\bar{a} - 0.195996\sigma, \bar{l} + 0.195996\sigma$	YES	0.391992 σ
$\bar{m}$	$\bar{m} - 0.195996\sigma, \bar{m} + 0.195996\sigma$	NO	0.391992 σ
$\bar{n}$	$\bar{n} - 0.195996\sigma, \bar{n} + 0.195996\sigma$	YES	0.391992 σ
$\bar{o}$	$\bar{o} - 0.195996\sigma, \bar{o} + 0.195996\sigma$	YES	0.391992 σ
$\bar{p}$	$\bar{p} - 0.195996\sigma, \bar{p} + 0.195996\sigma$	YES	0.391992 σ
$\bar{q}$	$\bar{q} - 0.195996\sigma, \bar{q} + 0.195996\sigma$	YES	0.391992 σ
$\bar{r}$	$\bar{r} - 0.195996\sigma, \bar{r} + 0.195996\sigma$	YES	0.391992 σ
$\bar{s}$	$\bar{s} - 0.195996\sigma, \bar{s} + 0.195996\sigma$	YES	0.391992 σ
$\bar{t}$	$\bar{t} - 0.195996\sigma, \bar{t} + 0.195996\sigma$	YES	0.391992 σ

Figure 2 95% *CI* for the mean of health-care professionals with n = 100

@ECRTD-UK: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

British Journal of Education
Vol.10., Issue 8, pp. 1-17, 2022
Print ISSN: 2054-6351(Print)
Online ISSN: 2054-636X (Online)



Question 2: A large bank is interested in determining the average amount of money its customers keep in interest- bearing accounts. A random sample of $n = 16$ customers produced a sample mean $\bar{x}$, with a standard deviation s. Assume that the amount of money kept in interest-bearing accounts by customers in the bank is normally distributed.
Construct a (i) 90% (ii) 95% *CIs* for the true amount of money kept in interest-bearing accounts by all customers of this bank.
(iii) Do the intervals really contain the population parameter μ?
(iv) Practically demonstrate the true meaning of the 90% and 95% confidence intervals?

Solution
L: Groups, which of the theorems apply to this question?
G6: The conditions for theorem 4 apply because in small sample situations ($n < 30$), we cannot replace σ with s, the quantity $\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$ does not produce values of the standard normal z under these conditions.
For a 90% confidence interval, $1 - \alpha = 0.90 \Longrightarrow \alpha = 0.10$, and so $\frac{\alpha}{2} = 0.05$. Table 2 is used to find the point $t_{0.05}$ for a t distribution with $\nu = n - 1 = 16 - 1 = 15$ degrees of freedom, this gives $t_{0.05} = 1.753$. By substituting $n = 16$, and $t_{0.05} = 1.753$ into the interval $\bar{a} - t_{0.05}\left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{x} + t_{0.05}\left(\frac{s}{\sqrt{n}}\right)$, gives $\bar{a} - 1.753\left(\frac{s}{\sqrt{16}}\right) < \mu < \bar{a} + 1.753\left(\frac{s}{\sqrt{16}}\right)$, it reduces to $\bar{a} - 0.43825s < \mu < \bar{a} - 0.43825s$. This means we are 90% confident that the interval from $\bar{a} - 0.43825s$ to $\bar{a} + 0.43825s$ contains the value of μ.
L: For the 90% confidence intervals, 10 samples of the salaries of 16 customers are randomly selected from the population. The mean of each of the 10 samples is calculated by using the formula $\bar{a} = \frac{\sum a}{n}$. The means of the 10 samples, with their corresponding confidence intervals are indicated in table 5. As with the previous question, 9 out of the 10 samples will contain the population parameter μ, whilst 1 out of the 10 samples will not. Table 6 indicates the sample means

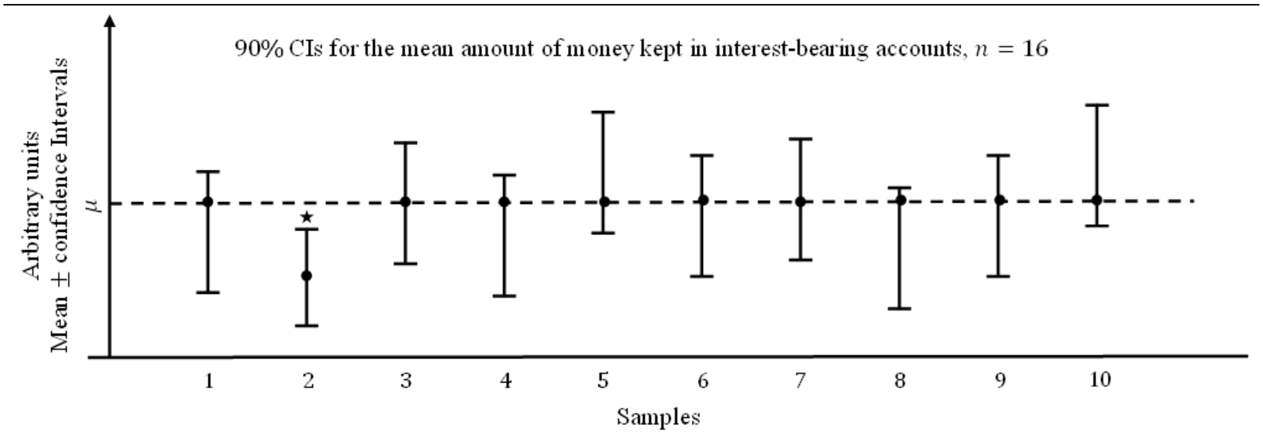
@ECRTD-UK: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

with 90% *CIs*, while Figure 3 indicates the 90% *CI* for the mean amount of money kept in interest-bearing accounts, *n* = 16

Table 6: Sample means with 90% *CIs*

Sample mean	90% *CIs*	Is the mean within the interval?	Width of the Interval
$\bar{a}$	$\bar{a} - 0.43825s, \bar{a} - 0.43825s$	YES	0.8765*s*
$\bar{b}$	$\bar{b} - 0.43825s, \bar{b} - 0.43825s$	NO	0.8765*s*
$\bar{c}$	$\bar{c} - 0.43825s, \bar{c} - 0.43825s$	YES	0.8765*s*
$\bar{d}$	$\bar{d} - 0.43825s, \bar{d} - 0.43825s$	YES	0.8765*s*
$\bar{e}$	$\bar{e} - 0.43825s, \bar{e} - 0.43825s$	YES	0.8765*s*
$\bar{f}$	$\bar{f} - 0.43825s, \bar{f} - 0.43825s$	YES	0.8765*s*
$\bar{g}$	$\bar{g} - 0.43825s, \bar{g} - 0.43825s$	YES	0.8765*s*
$\bar{h}$	$\bar{h} - 0.43825s, \bar{h} - 0.43825s$	YES	0.8765*s*
$\bar{\imath}$	$\bar{\imath} - 0.43825s, \bar{\imath} - 0.43825s$	YES	0.8765*s*
$\bar{\jmath}$	$\bar{\jmath} - 0.43825s, \bar{\jmath} - 0.43825s$	YES	0.8765*s*

Figure 3: 90% *CI* for the mean amount of money kept in interest-bearing accounts, *n* = 16

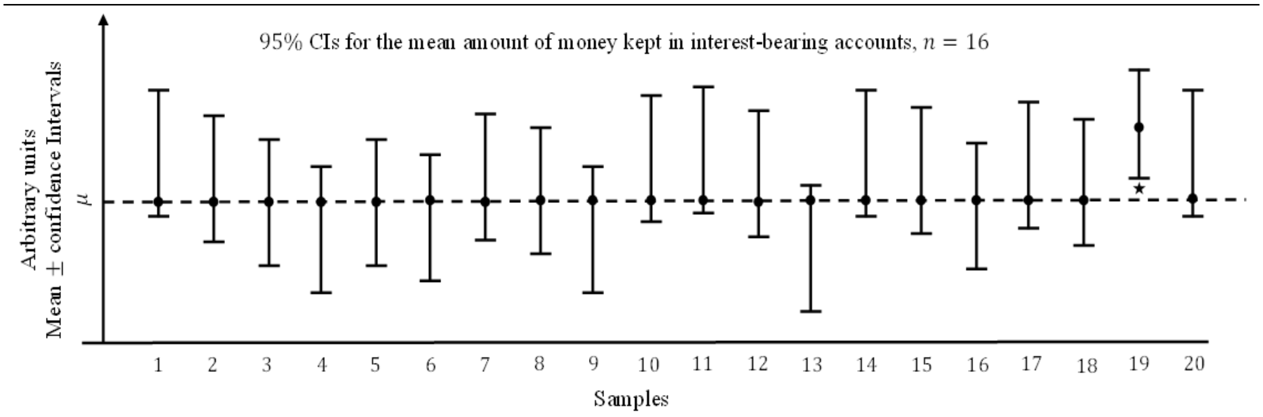


L: For the 95% *CIs*, 20 samples of the 16 customers were randomly selected from the population and their average amount of money was calculated using the same formula. An interval may be obtained as follows: $\bar{a} - (2.131)\left(\frac{s}{\sqrt{16}}\right) < \mu < \bar{a} + (2.131)\left(\frac{s}{\sqrt{16}}\right)$, which gives $\bar{a} - 0.53275s < \mu < \bar{a} + 0.53275s$. The means of the 20 samples, with their corresponding confidence intervals are indicated in table 6. Figure 4 indicates the 95% *CI* for the mean amount of money kept in interest-bearing accounts, *n* = 16

Table 6 Sample means with 95% *CIs*

Sample mean	95% *CIs*	Is the mean within the interval?	Width of the Interval
$\bar{a}$	$\bar{a} - 0.53275s, \bar{a} + 0.53275s$	YES	1.0654*s*
$\bar{b}$	$\bar{b} - 0.53275s, \bar{b} + 0.53275s$	YES	1.0654*s*
$\bar{c}$	$\bar{c} - 0.53275s, \bar{c} + 0.53275s$	YES	1.0654*s*
$\bar{d}$	$\bar{d} - 0.53275s, \bar{d} + 0.53275s$	YES	1.0654*s*
$\bar{e}$	$\bar{e} - 0.53275s, \bar{e} + 0.53275s$	YES	1.0654*s*
$\bar{f}$	$\bar{f} - 0.53275s, \bar{f} + 0.53275s$	YES	1.0654*s*
$\bar{g}$	$\bar{g} - 0.53275s, \bar{g} + 0.53275s$	YES	1.0654*s*
$\bar{h}$	$\bar{h} - 0.53275s, \bar{h} + 0.53275s$	YES	1.0654*s*
$\bar{\imath}$	$\bar{\imath} - 0.53275s, \bar{\imath} + 0.53275s$	YES	1.0654*s*
$\bar{\jmath}$	$\bar{\jmath} - 0.53275s, \bar{\jmath} + 0.53275s$	YES	1.0654*s*
$\bar{k}$	$\bar{k} - 0.53275s, \bar{k} + 0.53275s$	YES	1.0654*s*
$\bar{l}$	$\bar{l} - 0.53275s, \bar{l} + 0.53275s$	YES	1.0654*s*
$\bar{m}$	$\bar{m} - 0.53275s, \bar{m} + 0.53275s$	YES	1.0654*s*
$\bar{n}$	$\bar{n} - 0.53275s, \bar{n} + 0.53275s$	NO	1.0654*s*
$\bar{o}$	$\bar{o} - 0.53275s, \bar{o} + 0.53275s$	YES	1.0654*s*
$\bar{p}$	$\bar{p} - 0.53275s, \bar{p} + 0.53275s$	YES	1.0654*s*
$\bar{q}$	$\bar{q} - 0.53275s, \bar{q} + 0.53275s$	YES	1.0654*s*
$\bar{r}$	$\bar{r} - 0.53275s, \bar{r} + 0.53275s$	YES	1.0654*s*
$\bar{s}$	$\bar{s} - 0.53275s, \bar{s} + 0.53275s$	YES	1.0654*s*
$\bar{t}$	$\bar{t} - 0.53275s, \bar{t} + 0.53275s$	YES	1.0654*s*

Figure 4 95% *CI* for the mean amount of money kept in interest-bearing accounts, *n* = 16



Factors Affecting the Width of the CI

The width of the confidence interval is related to the confidence level, standard error, and *n* such that the following are true: The higher the percentage of confidence desired, the wider the confidence interval. The larger the standard error, the wider the confidence interval. The larger the *n*, the smaller the standard error, and so the narrower the confidence interval. All other things being equal, a smaller confidence interval is always more desirable than a larger one because a smaller interval means the population parameter can be estimated more accurately. A larger sample size expectedly will lead to a better estimate of the population parameter and this is reflected in a narrower *CI*. The width of the *CI* is thus inversely related to the sample size. In fact, the required sample size calculation for some statistical procedures is based on the acceptable width of the *CI*. Variability in a random sample directly influences the width of the *CI*. A larger spread implies that it is more difficult to reliably estimate population value without large amounts of data. Thus, as the variability in the data (often expressed as the *SD*) increases, the *CI* also widens.

RESULTS AND DISCUSSIONS

Majority of the students had various misconceptions about *CIs* despite having taken two introductory statistics courses at the university level. They may have developed more computational fluency than conceptual knowledge as a result of their lecturers' instructional strategies, which lend themselves to rote memorization rather than allowing the students to construct their own understanding. Another issue was that even though the lecturers may have used good teaching techniques, the students were unable to put the knowledge into practise. There are many misconceptions in the literature (Castro Sotos, Vanhoof, Van den Noortgate, & Onghena, 2007; Cumming & Maillardet, 2006; Fidler, 2006; Grant & Nathan, 2008; Greenland et al., 2016; Henriques, 2016). These misconceptions may have resulted from the inadequate teaching strategies used by lecturers.

When students are able to accurately define *CI* and acquire and appreciate its related theorems with the aid of their lecturers, their conceptual understanding of *CI* increases. This study has shown how crucial definitions and theorems are to students' conceptual understanding of *CI*. Before going on to other related ideas, lecturers should make sure that students fully understand these definitions and theorems.

Giving students advance notice of the guidelines and regulations that apply to tasks or assignments can help them maintain their focus on the tasks that their professors have given them. Before presenting an answer or idea as a group, students can actively participate in discussion when the social constructivism teaching method is used. With the main emphasis on students' conceptual understanding, teaching and learning in this manner heavily rely on interpersonal contact and conversation (Prawat, 1992).

Social constructivism improves knowledge about the process of knowledge formation (i.e., students determine as a group how they learn). It combines instruction with actual competitions (authentic tasks). It provides exposure and comprehension on many points of view (evaluations of alternative solutions). With the aid of constructivist teaching strategies, such as problem-solving and inquiry-based learning activities, students produce and test their ideas, draw conclusions and inferences, and pool and share their knowledge in a collaborative setting. Through observation, listening, and questioning, lecturers assist their students by using this technique.

This study has demonstrated that the width of a confidence interval largely depends on the confidence level, sample size, and standard deviation. The width of the confidence interval is directly proportional to the confidence level. Therefore, the larger the confidence level, the larger the width of that interval, and vice versa. Again, the width of the confidence interval is directly proportional to the standard deviation. The larger the standard deviation, the larger the width, and vice versa. The width of the confidence interval is also inversely proportional to the sample size. The larger the sample size, the smaller the width, and vice versa. It is also worth mentioning that with smaller sample sizes, the variability within the samples makes them attain large standard deviations, thereby having larger widths.

The intervals with asterisks are those which do not contain the population mean. For this study, they appear once at each confidence level. For the 90% and 95% confidence levels, the number of intervals is ten (10) and twenty (20), respectively. It needs emphasising in theory that, if a number of *CIs* were constructed at given levels of confidence, 90% and 95% of them would contain the population mean, while 10% and 5% would not.

Implications of this Study to Teaching and Learning

Lecturers should ensure that their instructional methods move away from rote learning toward conceptual learning. They should make students aware that conceptual learning could also be socially constructed. lecturers should use innovative instructional methods that collaboratively involve all students in the learning process. Such methods should evolve gradually from lecturers' own instructional practices for them to experiment many methods. In this regard, they should be motivated to participate in effective teacher professional development practices, to showcase or demonstrate their instructional practices. Eventually, effective instructional methods would be created that would better serve their students. These would require going past the conventional communication approach toward, including student involvement and sharing, hands-on exercises, and learning experiences. Conceptual learning would enable students to progress by encouraging understanding, encouraging dynamic learning systems, focusing on the result of the learning procedure, and relating new data to earlier learning.

Through conceptual learning, lecturers could encourage and motivate their students to use their prior experiences to understand new concepts. Conceptual learning could enable lecturers and students to understand how ideas relate to one another and create examples that will support them throughout their education and careers. By building on a solid foundation that supports a grasp of multiple ideas, conceptual learning could promote future learning. For example, a student who has a conceptual grasp of sampling techniques can work on confidence intervals with ease. Lecturers and students both have a significant role to play, the lecturer could provide tasks or assignments that motivate the students to fully develop an idea. Curriculum and instruction should be developed to facilitate the development of mental models or schema that link concepts essential to the subject or topic in order to support student learning. These objectives should be pursued in a methodical order using real-world activities and easily comprehensible illustrations that build on students' past knowledge and highlight the salient features of the material to be learnt.

CONCLUSION

By relying on concrete illustrations and examples, lecturers could help their students learn collaboratively to enhance their conceptual knowledge. Further, students' ability to understand definitions and apply them appropriately, could help them understand statistical concepts. This study, underpinned by the Vygotsky's social constructivism learning theory, allowed the establishment of opportunities for students to collaborate with their lecturers and peers in the construction of knowledge. The theory welcomed effective dialogue, collaboration, and application of information, as a significant aspect of learning among the students. As they interacted with their peers, they acquired the knowledge and skill to conceptually understand issues related to confidence interval of the population mean.

REFERENCES

American Psychological Association (2010). *Publication manual of the American psychological association* (6th Ed.). Author.

Bude, L., Imbos, T., van de Wiel, M. W. J., & Berger, M. P. (2011). The effect of distributed practice on students' conceptual understanding of statistics. *Higher Education*, *62*(1), 69–79.

Bude, L., van de Wiel, M. W. J., Imbos, T., & Berger, M. P. F. (2010). The effect of directive tutor guidance on students' conceptual understanding of statistics in problem-based learning. *British Journal of Educational Psychology*, *81*(2), 309–324.

Coulson, M., Healey, M., Fidler, F., & Cumming, G. (2010). Confidence intervals permit, but do not guarantee, better inference than statistical significance testing. *Frontiers in Psychology, 1*, 26.

Cumming, G. (2006). Understanding replication: Confidence intervals, p-values, and what's likely to happen next time. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics* (ICOTS-7). International Statistical Institute.

delMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. Statistics *Education Research Journal*, 62(2), 28–58.

Fisher, R.A. (1973). *Statistical methods for research workers* (14th ed.). Hafner Press.

Garfield, J., & Chance, B. (2000). Assessment in Statistics Education: Issues and Challenges. *Mathematics Thinking and Learning*, 2, 99-125.

Gosset (1908). The probable error of a mean. *Biometrika*, 6, 1–25

Graham, A. T., & Thomas, M. O. J. (2005). Representational versatility in learning statistics. *International Journal for Technology in Mathematics Education*, *12*(1), 3–4.

Hiebert, J., & Lefevre, P. (1986). Conceptual and Procedural Knowledge in Mathematics: An Introductory Analysis. In J. Hiebert (Ed.), *Conceptual and Procedural Knowledge: The Case of Mathematics* (pp. 1-27). Lawrence Erlbaum Associates.

Hong, E., & O'Neil Jr., H. F. (1992). Instructional strategies to help learners build relevant mental models in inferential statistics. *Journal of Educational Psychology*, *84*(2), 150–159.

Jones, J. L., Jones, K. A., & Vermette, P. J. (2011). Planning learning experiences in the inclusive classroom: Implementing the three core UDL principles to motivate, challenge and engage all learners, *Electronic Journal for Inclusive Education, 2* (7).

Koshy, V. (2010). *Action research for improving educational practice*. A step-by-step guide (2nd ed.). Sage.

Leppink, J., Broers, N. J., Imbos, T., van der Vleuten, C. P. M., & Berger, M. P. F. (2012). Self-explanation in the domain of statistics: An expertise reversal effect. *Higher Education*, 63, 771–785.

Lockwood, E., Yeo, A., Crooks, N. M., Nathan, M. J., & Alibali, M. W. (2014). Teaching about confidence intervals: How instructors connect ideas using speech and gesture. In W. Penuel, S. A. Jurow, & K. O'Connor (Eds.), *Learning and becoming in practice: Proceedings of the Eleventh International Conference of the Learning Sciences*. University of Colorado.

Meletiou-Mavrotheris, M., & Lee, C. (2002). Teaching students the stochastic nature of statistical concepts in an introductory statistics course. *Statistics Education Research Journal*, 1(2), 22–37

Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, *65*(2), 123–137.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skills in statistics: A cognitive load approach. *Journal of Educational Psychology*, *84*, 429–434.

Pfannkuch, M., Wild, C. J., & Parsonage, R. (2012). A conceptual pathway to confidence intervals. *ZDM– International Journal on Mathematics Education*, *44*(7), 899–911.

Smith, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, *3*(2), 115-163. https://doi.org/10.1207/s15327809jls0302_1