The Mathematics Enthusiast

Volume 16
Number 1 *Numbers 1, 2, & 3*

Article 24

2-2019

Pre-service teacher statistical misconceptions during teacher preparation program

Tracey Gorham Blanco

Scott A. Chamberlin

Pre-service teacher statistical misconceptions during teacher preparation program

Tracey Gorham Blanco
New Mexico State University
&
Scott A. Chamberlin[1]
University of Wyoming

Abstract: The purpose of the investigation was to identify pre-service elementary teachers' (PST) misconceptions at the culmination of their methods semester, prior to entering student teaching. Participants, *n*=134 (116 female) were pre-service elementary teachers from two universities in the intermountain region. The Statistical Reasoning Assessment (SRA) developed by Garfield (2003) was used to investigate student misconceptions in statistics and probability. Of the eight misconceptions, the Representativeness misconception and the Outcome orientation misconception were the least common (12.3 and 28.2% respectively) and the Comparing groups of the same size, Equi-probability bias, and Correlation implies causation misconceptions were the most common (70.2, 64.3, and 50.0% respectively). The confidence interval for the results was within a window of .389 to .427. Implications from the study are several, including a stand-alone statistics and probability course would likely improve PST's understanding of concepts in the domains, misconceptions should be used to promote true understanding, and preparers of PST should carefully analyze their students to gain legitimate understanding of their knowledge and misconceptions in statistics and probability.

Keywords: Statistics education; teaching and learning of probability; teaching and learning of statistics; misconceptions

The preparation of pre-service elementary teachers (PST) in mathematics is of paramount importance given two facts. First, increased attention to student performance on standardized tests has immediate effects on the teaching and learning process and second, the *Principles and Standards for School Mathematics* (NCTM, 2000) and the Common Core State Standards-Mathematics (NGA & CCSSO, 2010) expect conceptual understanding of domains in the content

[1] Scott@uwyo.edu

The Mathematics Enthusiast, **ISSN 1551-3440, vol. 16, nos.1,2&3**, pp. 461-484

area of mathematics. Regarding the *Principles and Standards for School Mathematics* (PSSM) document, statistics and probability is considered one of five content areas in K-12 and the Common Core State Standards-Mathematics (CCSSM) considers it one of several domains, based on the grade level. In short, statistics and probability has realized a considerable increase in importance in the past several decades (Shaughnessy, 2007) and can no longer be neglected by elementary schools and therefore it cannot be neglected by teacher preparation programs. The Guidelines for Assessment and Instruction in Statistics Education (GAISE) assert as much (Franklin, Kader, Mewborn, Moreno, Peck, Perry, & Scheaffer, 2007).

In this investigation, data are shared regarding PST misconceptions about statistics and probability using the Statistical Reasoning Assessment (SRA) that was created by Garfield at the turn of the millennium (Garfield, 2003). The instrument was designed to enable researchers to investigate correctness and student misconceptions. The focus of this article is on student misconceptions and there are eight included in this data set. Student (n=134) data was disaggregated and analyzed using various subgroups such as gender (females=116 and males=18), degree seeking status (traditional=87 and non-traditional=28) and previous statistics coursework (previous coursework=98 and no previous coursework=36). Data reveal several important findings. Following the review of extant literature, the method used to conduct the investigation is detailed, then the results are provided, with a subsequent discussion of results, including implications, limitations, and areas for future research.

Review of the Literature

General mathematics content of pre-service teachers

The connection between teacher preparation factors and student learning, particularly as measured by standardized assessments, is tenuous at best (Berliner, 2015). Simply stated, the case for teacher preparation factors having an immediate impact on student learning is not fully understood by researchers in the field of the psychology of mathematics education (nor is it precisely understood by researchers in any area of education). However, to suggest that teacher understanding, or lack thereof, has no part in facilitating student understanding of concepts in mathematics is naïve. After all, as Ma (1999) suggests, the manner in which teachers are prepared is instrumental in ultimately affecting change in and precipitating student understanding of concepts in mathematics. Knowing concepts deeply in mathematics is considered foremost in one's ability to direct learning environments (Ball & Bass, 2000). It is referred to as Mathematical Knowledge for Teaching or MKT. The intermediate factor in the equation, or one that links factor A (teacher performance) to factor C (student understanding) is factor B, which is teacher self-efficacy. According to Sutton and Krueger (2002) teachers without adequate content knowledge may be more inclined than their more informed peers to rely on rote memorization and textbook procedures to 'get by' in mathematics instruction. In short, the deeper one knows mathematical concepts, the more inclined that person is to teach well (Brown & Borko, 1992). The National Mathematics Advisory Panel (2008) substantiates this claim and further suggests that deep content knowledge implies the ability to make connections to grades before and after the intended curriculum.

Without identifying specific factors, Hill, Rowan, and Ball (2005) identified a significant relationship between what teachers know about mathematics (or content knowledge) and the direct application to student gains in achievement. In earlier work, Ball (1990) and more recently Silverman and Thompson (2008) mentioned the crucial nature of mathematics content and methods courses in relation to teachers' content knowledge and their understanding of concepts.

Thus, it *should be* the intent of all teacher preparation programs and affiliated individuals to engender deep conceptual understanding of domains in mathematics to such an extent that pre-service elementary teachers have some degree of comfort with facilitating learning episodes. Certainly statistics and probability is not the only content area or domain of mathematics. Nevertheless, it is an important one and thus true understanding is requisite in PST prior to entering the elementary classroom.

Student misconceptions of statistics and probability

One commonality permeates research of student conceptions and misconceptions in statistics. That commonality is that student understanding of concepts in statistics and probability is not impressive. As an example, in looking at National Assessment of Educational Progress (NAEP) data in statistics, it was apparent to Zawojewski and Shaughnessy (2000) as well as Tarr and Shaughnessy (2007) that students' understanding of statistics and probability is improving, but still not satisfactory. It may be argued that the improvement in test performance is simply a result of increased attention to the domain(s) of statistics and probability and not necessarily enhanced conceptual understanding. Nevertheless, Zawojewski, Shaughnessy, and Tarr admonish stakeholders that conceptual knowledge of statistics and probability is not likely deep. Shaughnessy (2007) in fact stated that much of the improvement is likely due to the fact that students started at such abysmally low levels that improvement was almost certain to occur.

Regarding the psychology of mathematics, the main topics that have been studied pertain to variability (Bakkar & Gravemeijer, 2004; Ciancetta, Shaughnessy, & Canada, 2003), average (Cai, 1995; Mokros & Russell, 1995), measures of center (Groth, 2005), inference (Hammerman & Rubin, 2003; Watson, 2001), manners in which graphical representations are used (Aberg-Bengtsson & Ottoson, 1995), and basic chance in probability (Chernoff & Sriraman, 2013;

Shaughnessy & Zawojewski, 1999; Sriraman & Chernoff, 2018). In studies relating to variability, Bakkar and Gravemeijer (2004) found that misconceptions involving variability are the biggest barrier to understanding data distributions and Ciancetta, Shaughnessy, and Canada (2003) showed that students generally view data sets as being drastically different because of variability even though such data sets have the same or similar measures of center. When studying averages and other measures of center, Cai (1995) found that middle school students' base knowledge of working with averages lies strictly in following the algorithm to find a mean and that it takes a considerable amount of conceptualization for such students' to gain a more complex understanding of the concept of an average. Similarly, Mokros and Russell (1995) analyzed student conceptualizations of averages and found that without specific instructional interventions, students generally use only the most basic concept for calculating an average and are not necessarily developing useful conceptions of an average. Groth (2005) also identified that students who reason about the measure of center in the context of the problem have more sophisticated levels of thinking than students who solve algorithmically without context. When reasoning about inference, Watson (2001) found that a visual representation of data allowed students to conceptualize variation, and make more appropriate decisions regarding the data, much more so than without such a visual. Consequently, student understanding of concepts in statistics and probability appears to be researched in much greater detail than teacher understanding of such concepts.

In-service teacher (mis)conceptions of statistics and probability

As early as 1988, Rubin and Roseberry found that when teachers investigate concepts as students do, there is ample room for improvement. The suggestion from their study was that additional efforts be invested in helping teachers understand concepts in statistics and probability, through the lens of a statistician. Fourteen years later, Makar and Confrey found similar results when they investigated teachers' understanding of variability in terms of their own students' high-stakes test data. Mikelson and Heaton conducted two separate studies (2004; 2003) with a single teacher-participant and found that the teacher's knowledge was suitable in some cases and incomplete in others. This teacher participant had specialized statistical professional development, but had a difficult time relating such professional development strategies to the content that was taught in the classroom. Therefore, the conceptual understanding of what was learned in the professional development setting was not translating to the setting of planning for teaching. Such studies supported findings with other research approaches and samples. More recently, Pfannkuch (2007) investigated teachers' understanding of box-plots and found not so much a lack of understanding of box plots, inasmuch as they found an inability of teachers to make sense of them enough to facilitate learning episodes. All of these studies support the claim made initially in 1993, by Shaughnessy and Bergman, that a significant chasm exists between what teachers are expected to teach and what they know.

Pre-service teacher (mis)conceptions of statistics and probability

Somewhat recently, an interest among researchers in statistics education has been pre-service teachers' (PST) understanding of concepts in statistics and probability. Groth and Bergner (2005) investigated PST knowledge of samples and how to sample, using metaphors for samples. Canada (2006) and Leavy (2006) respectively investigated PST understanding of variability and distributions. Leavy specifically found that statistical investigations conducted during a semester long methods course for PST improved their understanding of distributions by shifting the PSTs sole focus on descriptive statistics to graphical representations of data to improve understanding of such descriptive statistics relating to the distribution of data. Dollard

(2011) is credited with a strong study in which he investigated pre-service elementary teachers' misconceptions of probability using standard (e.g., a die) and non-standard shaped objects (such as a Monopoly® hotel) to see what their outcomes would be. One of the foci of his investigation was the concept of equiprobability and irregularity.

Given the recurring theme that understanding of concepts in statistics and probability is not at encouraging levels, the case for investigating PST misconceptions in the intermountain region is evident. In the next section, the methods employed to conduct the study, including the rationale for the instrument selection and participant selection, is made.

Method

Participant characteristics

In total, 137 participants were selected from a convenience sample of elementary PSTs at two universities in the Intermountain Region and 134 completed the demographic survey and the Statistical Reasoning Assessment [SRA] (Garfield, 2003) for a return rate of 98%. Participants' specific age was not recorded, but given the fact that all participants were going to student teach in the subsequent semester, it is assumed that the youngest was 21 years of age and the oldest participant cannot be determined. Of the 134 participants, 116 were female and 18 male. Ethnic and racial group affiliation was not recorded, though the majority of participants was of European-American heritage.

Sampling procedures

Participants were approached during the last weeks of their methods course (defined as courses designed to prepare student teachers to successfully deliver lessons in their student teaching semester), immediately prior to student teaching responsibilities. Initially, course coordinators were contacted to attain permission and then the first author distributed human consent forms, a demographic instrument (found in the appendix), and the SRA.

Measures and covariates

All participants were provided approximately one hour to complete the assessment that was comprised of 20 items. The SRA is comprised of items that enables researchers to investigate misconceptions as well as understandings in probability and statistics. In this article, the sole focus is on misconceptions, as detailing both understandings and misconceptions is beyond the scope of one article. The demographic instrument is found in the appendix and the SRA can be located in several publications, including Garfield's seminal 2003 publication in the *Statistics Education Research Journal*.

Research design

To duplicate the study, one need only access the SRA and the demographic survey contained in the appendix, identify a sample, implement the instruments, and interpret the data. To interpret the data, items were scored for correctness, using Garfield's key, and then data is reported as a percent correct. As an example, if the level of a misconception is .323, this data point indicates that less than one-third, more specifically 32.3%, of the participants selected an answer that was linked to the measured misconception. Naturally, the higher the score, the greater the number of participants revealed their misconception and the lower the score, the fewer number of participants that answered the item with the measured misconception. It is for this reason that true understanding is displayed by a data point that indicates a low degree of an identified misconception (e.g., 25% or lower).

Results

The results section is broken into two sections. In the first section, demographic data from participants is displayed. In the second section, results from the eight misconception scales is detailed, as per the demographic groups.

Demographic data

Demographic data is detailed in table 1. Prior to sharing the demographic data, a few definitions are requisite. For instance, student status refers to traditional or non-traditional students. Traditional students are those that plan on finishing their undergraduate degree, 5 years or less from their high school graduation. Non-traditional students, therefore, are those that will not complete their undergraduate degree within five years of high school graduation. Prior experience in a statistics course was defined as some coursework with a sole focus on statistics and/or probability. Finally, the degree sought was investigated and the two categories that comprised this factor were those seeking their first bachelor's degree and those seeking a second bachelor's degree and/or a master's degree in addition to those seeking a post-baccalaureate teaching certificate.

Table 1
Frequencies and Percentages for Demographic Variables

Demographic	**Category**	*n* (%)
Gender	Female	116 (86.6%)
	Male	18 (13.4%)
Student status*	Traditional	87 (64.9%)
	Non-traditional	38 (28.4%)
Previous statistics experience	Completed a stand-alone course	98 (73.1%)
	Did not complete a stand-alone course	36 (26.9%)
Type of degree sought	First bachelor's degree	111 (82.8%)
	Second bachelor's and/or master's degree	23 (17.2%)

9 participants did not respond to this item

Misconception data

In this section, data on misconceptions is provided relative to the four demographic categories presented. Commentary on the data is reserved for the discussion section. In table 2, the overall misconception constructs are presented, in which low numbers represent low prevalence of the misconception. As an example, it can be seen that the lowest prevalence of misconception was MC number 5 and the most prevalent misconceptions were MC's 7 and 8. It is also important to note that all *t*-tests conducted were Levene's Test for Equality of variances, which was used to determine the accurate statistical procedures for samples with assumed equal variances or those in which equal variances cannot be assumed.

Table 2: Aggregated scores of all participants by MC

Misconception Construct	*n*	M	CI	SD
MC1: averages	131	.337	.307-.368	.176
MC2: representation	132	.282	.245-.318	.211
MC3: good samples must be large	134	.366	.303-.429	.369
MC4: law of small numbers	133	.312	.264-.360	.279
MC5: representativeness	133	.123	.087-.159	.211
MC6: correlation implies causation	134	.500	.414-.586	.502
MC7: equiprobability bias	131	.643	.586-.700	.331
MC8: groups must have same *n* to be compared	134	.702	.623-.780	.459
MC total		.408	.389-.427	.109

Table 3: Misconception data disaggregated by gender

	Female				Male			
Construct	*n*	*M*	*SD*	*SE*	*n*	*M*	*SD*	*SE*
MC1	113	.352	.180	.017	18	.244	.110	.026
MC2	115	.294	.217	.020	17	.200	.141	.034
MC3	116	.375	.378	.035	18	.306	.304	.072
MC4	115	.309	.286	.027	18	.333	.243	.057
MC5	116	.118	.212	.020	17	.157	.208	.051
MC6	116	.517	.502	.047	18	.389	.502	.118

MC7	114	.634	.333	.031	17	.706	.321	.078
MC8	116	.724	.449	.042	18	.556	.511	.121
Misconception Score	112	.413	.110	.010	17	.373	.096	.023

Table 4: *t*-test and confidence interval for means by gender

	t-test for variances					95% CI of the Difference	
Construct	*F*	Sig	*t*	df	Sig. (2-tailed)	Lower	Upper
MC1	4.925	.028**	-3.489	33.7	.001**	-.1706	-.0450
MC2	8.393	.004**	-2.357	28.6	.025**	-.1754	-.0124
MC3	1.710	.193	-.742	132	.459	-.2545	.1156
MC4	2.355	.127	.347	131	.729	-.1159	.1652
MC5	.150	.699	.711	131	.478	-.0696	.1477
MC6	4.861	.029**	-1.010	22.6	.323	-.3915	.1348
MC7	2.026	.157	.837	129	.404	-.0984	.2426
MC8	3.882	.051	-1.46	132	.148	-.3978	.0606
Misconception Score	.864	.354	-1.43	127	.156	-.0963	.0156

Table 5: Misconception data disaggregated by student status

	Traditional				Non-Traditional			
Construct	*n*	*M*	*SD*	*SE*	*n*	*M*	*SD*	*SE*
MC1	85	.341	.177	.019	38	.305	.166	.027
MC2	87	.306	.230	.025	37	.216	.166	.027
MC3	87	.397	.390	.042	38	.276	.323	.052
MC4	87	.328	.284	.030	38	.303	.274	.044
MC5	87	.123	.210	.023	37	.126	.213	.035
MC6	87	.494	.503	.054	38	.500	.507	.082
MC7	87	.644	.334	.036	36	.660	.323	.054
MC8	87	.747	.437	.047	38	.579	.500	.081
Misconception Score	85	.425	.110	.012	36	.369	.103	.017

Table 6: *t*-test for Equality of Means for Student Status

Construct	*t*-test for variances					95% Confidence Interval of the Difference	
	F	Sig.	*t*	df	Sig. (2-tailed)	Lower	Upper
MC1	.013	.910	-1.060	121	.291	-.1030	.0311
MC2	8.778	.004**	-2.436	92.9	.017**	-.1625	-.0165
MC3	1.844	.177	-1.667	123	.098	-.2630	.0225
MC4	.003	.956	-.457	123	.648	-.1330	.0830
MC5	.012	.912	.085	122	.932	-.0785	.0855
MC6	.005	.944	.059	123	.953	-.1883	.1998
MC7	.491	.485	.245	121	.807	-.1137	.1457
MC8	9.216	.003**	-1.794	62.8	.078	-.3555	.0191
Misconception Score	.062	.804	-2.602	119	.010**	-.0984	-.0133

Table 7: Group Statistics, Type of Degree

Construct	Undergraduate				Post-Bacc/Master's			
	n	*M*	*SD*	*SE*	*n*	*M*	*SD*	*SE*
MC1	109	.350	.181	.017	22	.273	.132	.028
MC2	111	.299	.216	.021	21	.191	.161	.035
MC3	111	.383	.387	.037	23	.283	.253	.053
MC4	111	.311	.287	.027	22	.318	.246	.053
MC5	111	.126	.211	.020	22	.106	.215	.046
MC6	111	.496	.502	.048	23	.522	.511	.107
MC7	110	.646	.336	.032	21	.631	.312	.068
MC8	111	.730	.446	.042	23	.565	.507	.106
Misconception Score	108	.417	.110	.011	21	.361	.093	.020

Table 8: t-test for Equality of Means for previous statistics experience beyond the math content course required for pre-service elementary teachers

Constructs	t-test for variances F	Sig.	t	df	Sig. (2-tailed)	95% Confidence Interval of the Difference Lower	Upper
MC1	.757	.386	-.329	129	.743	-.0710	.0508
MC2	.823	.366	.232	130	.817	-.0644	.0815
MC3	1.496	.224	-.467	132	.641	-.1562	.0965
MC4	.176	.675	-.990	131	.324	-.1438	.0479
MC5	.008	.929	-.086	131	.931	-.0759	.0695
MC6	.000	1.000	-.860	132	.391	-.2463	.0971
MC7	4.796	.030**	.312	126.2	.756	-.0964	.1324
MC8	8.977	.003**	-1.512	130.2	.133	-.2757	.0368
Mis-conception Score	2.539	.114	-2.087	127	.039**	-.0771	-.0020

Table 9: *t*-test for Equality of Means for Type of Degree

Construct	*t*-test for variances F	Sig.	t	df	Sig. (2-tailed)	95% Confidence Interval of the Difference Lower	Upper
MC1	2.123	.147	1.914	129	.058	-.002635	.158098
MC2	9.566	.002**	2.672	35.137	.011**	.026120	.191126
MC3	5.612	.019**	1.558	46.241	.126	-.029256	.229804
MC4	1.699	.195	-.113	131	.910	-.136823	.122081
MC5	.238	.626	.406	131	.685	-.077709	.117840
MC6	.157	.692	-.227	132	.820	-.254501	.202014
MC7	.585	.446	.183	129	.855	-.142096	.171100
MC8	5.020	.027**	1.445	29.477	.159	-.068186	.397211
Misconception Score	.679	.412	2.185	127	.031**	.0052952	.1067253

Discussion

Naturally, all of the data cannot be discussed in this section, so statistics of note are highlighted with some commentary. First, it is important to reiterate that a low score for misconceptions is a desirable statistic. That is to say, if a very low number of respondents hold a misconception, this practically means that few of the pre-service teachers (PST) in the sample have the misconceptions. Moreover, the fewer misconceptions, the better the teacher's knowledge.

Based on the premise that it is somewhat unnatural for all teachers to have no misconceptions, there must be a level of misconceptions that is *acceptable* per misconception. In looking at the lowest or least frequent misconception from the aggregated data in table 2, it is apparent that several misconceptions appear to be infrequent. In specific, the least common misconception (MC5) was representativeness (12.3% of individuals), which means that people estimate the likelihood of an event based on how closely the sample aligns with the population (Garfield, 2002). For example, if an individual rolls a die 17 times and has not gotten a value of six, then an individual with the representativeness misconception may be inclined to think that the six must come up on the next roll to mimic the actual representation of all numbers. Individuals without the representativeness misconception realize that the likelihood of a six on the 18^{th} roll is still 1 in 6. Also of note as somewhat infrequent misconceptions are the representation (MC2) in which 28.2% and the law of small numbers misconception (MC4) in which 31.2% of the pre-service teachers held this misconception. The representation misconception (Garfield) is one in which choosing a sample that is representative of the population, say for instance a truly random sample, is considered to be a better practice than choosing an abnormally large one because there is no guarantee that the large sample represents the population well. The law of small numbers misconception (Garfield) is one in which an abnormally small *n* in a sample might misrepresent the entire population, even if it was identified randomly.

A more troubling note were the misconceptions that occurred at a very high frequency. While an agreed upon level of misconceptions will perhaps never be reached, certainly any misconception that is held by at least 50% of the PST that comprised the sample is of note. The least frequently occurring misconception (in the high category) was MC6 (Garfield, 2002), that correlation implies causation and 50% of the participants held this misconception. To use an example, if height and the propensity to be in the National Basketball Association (NBA) correlate, then those with this misconception feel that height necessarily caused membership in the NBA. Also, MC7 (Garfield) which is known as the equiprobability misconception, occurred in 64.3% of the PST. Simply stated, the equiprobability bias occurs when individuals misinterpret the likelihood of an event transpiring as equal, when two events are not equally likely. For instance, if 12 doctors and 4 nurses were in a meeting, individuals with the equiprobability bias might misinterpret the likelihood of selecting a nurse randomly from the sample as the same probability as selecting a doctor. Finally, and perhaps most disconcerting, was the fact that 70.2% of the PST in this sample held MC8 (Garfield) that refers to the notion that groups must have the same number of events or people in the sample for adequate comparison. If this were true, then it would not be safe for a statistician to compare two groups, one comprised of 738 teachers and another of 268 administrators, to be compared. In any event, the frequency of some of the misconceptions almost certainly guarantees that future teachers will not be able to adequately teach their students about relatively basic concepts in statistics and probability.

Despite the most common misconception (MC8), that groups of different size cannot be compared, it is arguably problematic to compare two groups of the disparity with the disparity of males and females as found in table 3, because it cannot be assumed that the group of males was large enough to have a normal distribution, thus perhaps requiring the use of non-parametric statistics. Nevertheless, the data is provided for perusal and it simply suggests that males did realize a slightly lower overall level of misconception (.373) than females did (.413). As the law of small numbers is applied appropriately, it is important to note that with such a small sample of males, it is possible that 4-5 highly astute men in probability and statistics may have ballooned the mean for males, while the females had regression towards the mean with a sample of approximately 120 individuals. The only statistical differences at the .05 level are MC1 (averages), MC2 (representation), and MC6 (correlation implies causation), as found in table 4, but again these data must be interpreted with extreme caution given the low *n* in the male sample.

In tables 5 and 6, data are presented about misconceptions by student status (traditional versus non-traditional). To revisit the definition, traditional students were ones that were on track to graduate five years or less from their high school graduation date and non-traditional students were those that would require more than five years after high school to graduate. This was the only metric used to define student status. As table 6 indicates, the only statistical difference in the two groups at the .05 level was on MC2 (representation) and MC8 (groups must have the same *n* to be compared). Also, the data was disaggregated by type of degree sought using bachelor’s degree as one category and post-baccalaureate and master’s degree combined as the second category. The only misconception that appeared as a statistical difference at the .05 level was MC2 (representation). When previous statistical experience was analyzed, there were no statistical differences found for any of the individual constructs, however a statistical difference at the .05 level was found for the overall misconception scores for such experience. Finally, a quick overview of the data was performed to identify the lowest and highest level of misconception by disaggregated data. The lowest level of misconception was attained by females on MC5 (representativeness) in which 11.8% of the females that took the SRA held the misconception. That translates to slightly less than 1 in 8 females with the representativeness misconception. The highest level of misconception was held by traditional students on MC8 (groups must have the same *n* to be compared), which occurred at a frequency of 74.7% or practically speaking 3 out of 4 traditional students assumed that groups must be of the same size for comparison.

Implications

Two implications came about as a result of the study. First, if this sample is indicative of pre-service teachers' misconceptions in statistics and probability as a whole, the three most common misconceptions (MC6, MC7, and MC8) must be addressed in teacher content courses embedded in teacher preparation programs. Even the lesser occurring misconceptions (e.g., 1, 2, 3, and 4) are rather high. In fact, given some of the levels of misconception, it may be argued that programs that question the statistics and probability understanding of their candidates should strongly consider gathering data with the prospective objective of increasing exposure and experiences in the two interrelated domains.

Second, the eight misconceptions that Garfield and colleagues have identified in the almost 20 years since the initial SRA was designed in 1998, must not be the only ones apparent with elementary PSTs. Teacher preparation programs must consider expanding the number of misconceptions to those beyond the eight presented, using something such as the Guidelines for Assessment and Instruction in Statistics Education (GAISE) report (Franklin, et al., 2007). Assuming this sample is representative of most PST, the level of understanding with some concepts is alarming.

Limitations

All studies have some limitations and the primary limitation with this study was the low *n* of participants. This issue was exacerbated by the few number of teacher preparation programs involved (2). That is to say, if the number of teacher preparation programs and the number of participants in the overall sample would have been expanded to a greater number, researchers may have been more confident in making generalizations about the findings. A second limitation is that though the SRA is a well-respected instrument but there may be additional misconceptions worthy of analysis. Hence, looking for additional teacher misconceptions, maybe initially through qualitative approaches, would enhance subsequent findings. Estrella, Olfos, and Mena-Lorca (2015) did create a separate instrument that has additional concepts on it.

Areas for future research

The opportunity to ameliorate shortcomings in research, so-called limitations, comes in the form of future research. That is to say, the best research is that which has helpful findings, but which helps researchers realize that limitations must be addressed in future iterations of data collection. To that end, securing a much larger *n* would be instrumental in generalizing results. Possibly before identifying a larger *n*, researchers should investigate statistics and probability misconceptions with qualitative approaches, with the intent of piloting those items on an addition to the SRA.

Conclusion

The future of success in elementary mathematics teaching does not hinge on the outcome of this study, but the data does precipitate some concerns about PST understanding, or lack thereof, regarding somewhat basic concepts in statistics and probability. For instance, any time that the level of misunderstanding of concepts equals or exceeds 50%, as was the case with three of the misconceptions investigated with the use of this instrument individuals preparing teachers should have concern. The sample was not abnormally large, but it also was not particularly small (approaching an *n* of 150). Hence, a consistent pattern has emerged as areas of concern. There were some positive notes from the research. For instance, the frequency of misconception with at least one of the eight misconceptions, namely MC8 (representativeness), was encouraging. It might be assumed that with a modicum of review, this misconception could be all but eliminated by participants in this sample.

References

Aberg-Bengtsson, L. & Ottoson, T. (1995). *Children's understanding of graphically represented quantitative information.* Paper presented at the sixth Conference of the European Association for Research on Learning and Instruction. Nijmegen, The Netherlands.

Bakkar, A. & Gravemeijer, K. (2004). Learning to reason about distribution. In J. Garfield & D. Ben-Zvi (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147-168). Dordrecht, The Netherlands: Kluwer.

Ball, D. (1990). Prospective elementary and secondary teachers‟ understanding of division. *Journal of Research in Mathematics Education, 21*, 132-144.

Ball, D. & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed)., *Multiple perspectives on the teaching and learning of mathematics,* (pp. 83 – 104). Westport, CT: Ablex.

Berliner, D. (2015). Teacher evaluation and standardized tests: A policy fiasco. Video retrieved on 7 January 2016 from http://education.unimelb.edu.au/news_and_activities/events/upcoming_events/dean_lecture_series/dls-past-2015/teacher-evaluation-and-standardised-tests-a-policy-fiasco

Brown, C. & Borko, H. (1992). Becoming a mathematics teacher. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning.* (pp. 209-242). NewYork: Macmillan.

Cai, J. (1995). Beyond the computational algorithm: Students' understanding of the arithmetic average concept. In L. Meira & D. Carraher (Eds.), *Proceedings of the 19th Psychology of Mathematics Education Conference* (Vol. 3. Pp. 144-151). Sao Paulo, Brazil: PME Program Committee.

Canada, D. (2006). Elementary pre-service teachers' conceptions of variation in a probability context. *Statistics Education Research Journal, 5*, 36-63.

Chernoff E.J., & Sriraman,B. (2013). *Probabilistic Thinking: presenting plural perspectives*. Springer Berlin.

Ciancetta, M., Shaughnessy, J., & Canada, D. (2003). Middle school students' emerging definitions of variability. In N. Pateman, B. Dougherty, & J. Zilliox (Eds.), Poster Session in the *Proceedings of the 27th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, p. 481). Honolulu, HI: University of Hawaii.

Dollard, C. (2011). Pre-service elementary teachers and the fundamentals of probability. *Statistics Education Research Journal, 10*, 27-47.

Estrella, S., Olfos, R., & Mena-Lorca, A. (2015). Pedagogical content knowledge of statistics among primary school teachers. *Educação E Pesquisa, 41*, 477-492.

Franklin, C., Kader, G., Mewborn, D., Moreno, J. Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K – 12 curriculum framework.* Alexandria, VA: American Statistical Association.

Garfield, J. (2003). Assessing statistical reasoning. *Statistics Education Research Journal, 2,* 22-32.

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, *10*, http://www.amstat.org/publications/jse/v10n3/garfield.html

Groth, R.E. (2005). An investigation of statistical thinking in two different contexts: Detecting a signal in a noisy process and determining a typical value. *Journal of Mathematical Behavior, 24*, 109-124.

Groth, R. & Bergner, J. (2005). Pre-service elementary teachers' metaphors for the concept of statistical sample. *Statistics Education Research Journal*, *4*, 27-42.

Hammerman, J. & Rubin, A. (2003). Reasoning in the presence of variability. In C. Lee (Ed.), *Proceedings of the Third International Research Forum on Statistical Reasoning, Thinking, and Literacy*. Mt. Pleasant: Central Michigan University.

Hill, H.C., Rowan, B., & Ball, D. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*, 371- 406.

Leavy, A. (2006). Using data comparisons to support a focus on distribution: Examining pre-service teachers' understanding of distribution when engaged in statistical inquiry. *Statistics Education Research Journal, 5*, 89-114.

Ma, L. (1999). Knowing and teaching elementary mathematics. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Makar, K., & Confrey, J. (2002). *Comparing two distributions: Investigating secondary teachers' statistical thinking.* Paper presented at the Sixth International Conference on Teaching Statistics. Cape Town, South Africa.

Mikelson, W., & Heaton, R. (2004). Primary teachers' statistical reasoning about data. In D.

Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 327-352). Dordrecht, The Netherlands: Kluwer.

Mikelson, W. & Heaton, R. (2003). Purposeful statistical investigation merged with K-6 content: Variability, learning, and teacher knowledge use in teaching. In C. Lee (Ed.), *Proceedings of the Third International Research forum on statistical reasoning, thinking, and literacy*. Mt. Pleasant: Central Michigan University.

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, *26*, 20-39.

National Council of Teacher of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: National Council of Teachers of Mathematics.

National Governors Association & Council of Chief State School Officers. (2010). *Common core state standards for mathematics.* Washington, D.C.: National Governors Association Center for Best Practices. retrieved from http://www.corestandards.org/

National Mathematics Advisory Panel. (2008). Foundations of success: Final report of the National Mathematics Advisory Panel. Retrieved from http://www.ed.gov/ about/bdscomm/list/mathpanel/report/final-report.pdf

Pfannkuch, M. (2007). Comparing box plot distributions: A teacher's reasoning. *Statistics Education Research Journal*, *5*, 27-45.

Rubin, A., & Rosebery, A. (1988). *Teachers' misunderstandings in statistical reasoning: Evidence from a field test of innovative materials.* Paper presented at the International Statistics Institute Round Table Conference, Budapest, Hungary: International Statistics Institute.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957-1009). Reston, VA: National Council of Teachers of Mathematics.

Shaughnessy, J., & Bergman, B. (1993). Thinking about uncertainty: Probability and Statistics. In P. Wilson (Ed.) *Research ideas for the classroom: High school mathematics* (pp. 177-197). Reston, VA: National Council of Teachers of Mathematics.

Shaughnessy, J. & Zawojewski, J. (1999). Secondary students' performance on data and chance in the 1996 NAEP. *The Mathematics Teacher, 92,* 713-718.

Silverman, J., & Thompson, P. W. (2008). Toward a framework for the development of mathematical knowledge for teaching. *Journal for Mathematics Teacher Education, 11,* 499 – 511.

Sriraman B., Chernoff E.J. (2018) Probabilistic and Statistical Thinking. In: Lerman S. (eds) *Encyclopedia of Mathematics Education*. Springer, Cham

Sutton, J., & Krueger, A. (Eds.). (2002). *EDThoughts: What we know about mathematics teaching and learning*. Aurora, CO: Mid-continent Research for Education and Learning.

Tarr, J., & Shaughnessy, J. (2007). Student performance in data analysis, statistics, and probability. In P. Kloosterman & F. K. Lester, Jr. (Eds.), *Results and interpretations of the 2003 mathematics assessment of the National Assessment of Educational Progress* (pp. 139-168). Reston, VA: National Council of Teachers of Mathematics.

Watson, J. (2001). Longitudinal development of inferential reasoning by school students. *Educational Studies in Mathematics, 47,* 337-372.

Zawojewski, J., & Shaughnessy, J. (2000). Data and chance. In E. A. Silver & P. A. Kenney

(Eds.), *Results from the seventh mathematics assessment of the National Assessment of Educational Progress* (pp. 235-268). Reston, VA: National Council of Teachers of Mathematics.