

METHODOLOGY

Quantitative Methodology: A Guide for Emerging Physical Education and Adapted Physical Education Researchers

Justin A. Haegele and Samuel R. Hodge

Abstract

Emerging professionals, particularly senior-level undergraduate and graduate students in kinesiology who have an interest in physical education for individuals with and without disabilities, should understand the basic assumptions of the quantitative research paradigm. Knowledge of basic assumptions is critical for conducting, analyzing, and presenting research of high quality in this arena. In this tutorial paper, we present information essential to understanding the assumptions undergirding the quantitative research paradigm including the logic of hypothesis testing and sampling. Moreover, we describe key aspects of true and quasi-experimental research designs commonly used in quantitative studies.

Research paradigms are undergirded by a set of assumptions (e.g., epistemology and ontology) that influence researchers' decisions and actions. Assumptions can be defined as a set of beliefs that guide the way in which researchers approach their investigations (Fraenkel, Wallen, & Hyun, 2012). Fraenkel et al. (2012) described assumptions as being "related to the views they [researchers] hold concerning the nature of reality, the relationship of the research to

Justin A. Haegele is an assistant professor, Department of Human Movement Sciences, Old Dominion University. Samuel R. Hodge is a professor, Department of Human Sciences, The Ohio State University. Please send author correspondence to jhaegele@odu.edu

that which he or she is studying, the role of values in a study, and the process of research itself” (p. 427). Assumptions then guide the research endeavor, including the methods used and the questions asked (Hathaway, 1995). Emerging professionals such as upper level undergraduate and graduate students interested in research pertaining to physical education for individuals with and without disabilities should understand the basic assumptions of the quantitative paradigm for conducting, analyzing, and presenting research of high quality. This paper will serve as a convenient resource that can be distributed to emerging professionals to provide introductory research content. It is a user-friendly guide to unlock the mystery of many research methods textbooks and courses. Therefore, there were two purposes of this paper. The first purpose was to provide readers with information essential to understanding the assumptions undergirding the quantitative research paradigm including the logic of hypothesis testing and sampling. The second purpose was to describe key aspects of true and quasi-experimental research designs.

Basic Assumptions of the Quantitative Research Paradigm

The basic assumptions for each research paradigm are related to the philosophy under which the paradigm is situated. The quantitative research paradigm is based on the philosophy of positivism. *Positivism* is supported by an external realist ontology, in which it is assumed that a hard reality exists (Fraenkel et al., 2012; Pringle, 2000). The philosophy of positivism has influenced the quantitative research paradigm by providing several assumptions that guide researchers’ actions. According to Fraenkel et al. (2012), there are eight major assumptions of quantitative research (see Table 1). Researchers must consider each assumption when designing and implementing research using the quantitative paradigm.

Table 1
Major Assumptions of Quantitative Research

1. A hard reality exists, and it is the task of science to discover the nature of reality and how it works.
2. Research investigations can result in accurate statements about the way the world really is.
3. Researchers may remove themselves from what is being researched.
4. Facts are independent of the knower (the person with the knowledge) and can be known in an unbiased way.

Table 1 (cont.)

-
5. Facts and values are distinct from one another.
 6. Proper research designs can lead to accurate conclusions about the nature of the world.
 7. The purpose of research is to explain and predict relationships.
 8. The goal of research is to develop laws that make prediction possible.
-

Note. Derived from *How to Design and Evaluate Research in Education* (8th ed.), by J. R. Fraenkel, N. E. Wallen, and H. H. Hyun, 2012, New York, NY: McGraw-Hill.

In basic terms, a positivist's view of the role of science is to discover the nature of this reality and determine how it works (Fraenkel et al., 2012). Furthermore, a positivist epistemology claims to be free of value and not influenced by social context (Pringle, 2000). Positivists believe their understanding of knowledge can be generalized to other individuals across different environments and time. Positivism states the world is deterministic, meaning all events that occur in the world are a result of a cause-and-effect relationship (Trochim & Donnelly, 2006). An essential aspect of this paradigm is the logic of hypothesis testing.

The Logic of Hypothesis Testing

Hypothesis testing is the formal procedure used by researchers utilizing the group design research paradigm to accept or reject statistical hypotheses. Hypothesis testing is based on the mathematic concept of probability, which represents the likelihood of an event occurring. A research hypothesis is the predicted outcome or the expected results from a study (Fraenkel et al., 2012; Thomas, Nelson, & Silverman, 2005). This anticipated result may be derived from previous literature, a theoretical framework, or a researcher's previous experiences (Gay, Mills, & Airasian, 2006; Thomas et al., 2005). The formulation of a research hypothesis is critical to the beginning of the research process, as every aspect of a study is affected by the hypothesis including the participants selected, design of the study, and data analysis strategy (Gay et al., 2006). In contrast to the research hypothesis, the null hypothesis states there was no change in the participants' behavior after the researcher introduced the intervention (Fraenkel et al., 2012). In null hypothesis testing, the researcher uses deductive reasoning to ensure the truth of conclusions is irrefutable (Wilkinson, 2013). A critical feature of a re-

search hypothesis and null hypothesis is they must be testable, providing a way the claim can be either supported or refuted (Thomas et al., 2005).

To further illustrate the difference between research and null hypotheses, consider a study by Cervantes and Porretta (2013), who sought to determine the effect of a social cognitive theory–based after-school program on the leisure-time physical activity participation of adolescents with visual impairments. For this study, a research hypothesis may state the expected results of the study would be that adolescents with visual impairments exhibit more leisure-time physical activity participation when introduced to a social cognitive theory–based after-school program. On the contrary, the null hypothesis would indicate adolescents with visual impairments do not exhibit a change in leisure-time physical activity participation when introduced to a social cognitive theory–based after-school program.

Several misconceptions exist about hypothesis testing (Thomas et al., 2005). One common misconception about hypothesis testing is that researchers seek to determine whether a research hypothesis is accurate. Statistics cannot prove a research hypothesis is correct (Thomas et al., 2005). Rather, researchers seek to determine whether the hypothesis is supported by data (Gay et al., 2006). Another common misconception is that researchers test their research hypothesis. Instead, it is always the null hypothesis that is tested. All that statistics do is inform the researcher to either reject or fail to reject the null hypothesis (Thomas et al., 2005).

After the *research* and *null hypotheses* are stated, the researcher will then test the null hypothesis to assess the probability of the sample result if the null hypothesis were true (Fraenkel et al., 2012). In other words, the researcher looks to determine the probability of whether the data obtained through the study could exist if the null hypothesis were true. To do so, researchers decide what statistics are relevant for the particular data set and specific hypothesis. Using the statistical procedure(s) chosen, researchers analyze the data to determine the probability of obtaining the sample results if the null hypothesis were true. The last step in hypothesis testing is to either accept or reject the null hypothesis. Researchers base this decision on the level of significance of the result. If the probability of obtaining the sample results with the null hypothesis being true is small, the null hypothesis is rejected. If the probability of obtaining the sample results with the null hypothesis being true is large, the

research hypothesis is not supported. But what is a large or small probability?

In most quantitative studies, including group designs commonly used in physical education research, a standard significance level is .05. For example, Vigo-Valentín, Bush, and Hodge (2014) examined the physical activity behaviors of 637 adolescents attending middle and high schools in Puerto Rico. They set the significance (called *alpha*, α) level a priori at 0.05 and reported, among other findings, that middle school students participated significantly, $F(1) = 4.73$, $p = .03$, more in moderate physical activities than did high school students. For this finding, the researchers would reject the null hypothesis of no difference between middle and high school students' physical activity participation because the p value was below .05. When researchers reject a null hypothesis at the .05 significance level, they are saying the probability of obtaining that outcome with the null hypothesis being true is only 5% (Fraenkel et al., 2012). This also means the probability of obtaining that outcome with the null hypothesis being false (or the research hypothesis being supported) is 95%. In most cases, a 95% probability of a research hypothesis being supported is acceptable and appropriate. However, this opens the door for researchers to make errors, as 95% does not account for all possibility. One error, called a type I error, occurs if a researcher rejects a null hypothesis that is true, or a false positive. This happens when the 5% chance of being correct is, in fact, correct. A second error is called a type II error, the magnitude of which is determined by *beta* (β ; Thomas et al., 2005). This error represents when a researcher fails to reject a null hypothesis that is false. A type II error is also known as a false negative and can result from a researcher who mistakenly believes there was no difference between treatment groups ($p < .05$), even though the difference was statistically significant. Beta error (or type II error) indicates the researcher's hypothesis was likely true, despite the decision to reject it and accept the null hypothesis (Levin, 1983). Thomas et al. (2005) explained the relationship between alpha and beta, in that "as alpha is set increasingly smaller, beta becomes larger" (p. 115). Figure 1 shows the four possibilities for hypothesis testing and error.

	Null hypothesis true	Null hypothesis false
Accepting	Correct	Type II Error
Rejecting	Type I Error	Correct

Figure 1. Hypothesis decision error table. Adapted from *How to Design and Evaluate Research in Education* (8th ed.), by J. R. Fraenkel, N. E. Wallen, and H. H. Hyun, 2012, New York, NY: McGraw-Hill, p. 233. Copyright 2012 by McGraw-Hill.

The logic of hypothesis testing can be summarized in six steps (Fraenkel et al., 2012). The sequence entails (1) stating a research hypothesis; (2) stating a null hypothesis; (3) determining the statistics that are pertinent to the hypothesis and type(s) of data (e.g., ordinal, interval); (4) determining the probability of obtaining the sample results if the null hypothesis is true; (5) rejecting the null hypothesis if the probability is small, hence confirming the research hypothesis; and (6) not rejecting the null hypothesis if the probability is large, hence not affirming the research hypothesis (Fraenkel et al., 2012). A visual display of this sequence is provided in Figure 2. Hypothesis testing requires the random or nonrandom selection of one or more samples for making inferences (i.e., statistical generalizations) about the target population. For example, Hodge, Davis, Woodard, and Sherrill (2002) used hypothesis testing to compare the effects of two practicum types on physical education teacher education students' attitudes and perceived competence toward teaching students with disabilities. The two samples were students enrolled in either Group A (off-campus practicums) or Group B (on-campus practicums).

Sampling Strategies

A *sample* is a group of participants on which a study is conducted (Thomas et al., 2005). The larger group of people whom the researcher hopes to infer the findings from the study is referred to as the population. One of the most critical elements of a study is selecting the individuals who will participate (Fraenkel et al., 2012). The process of selecting participants for a research study from the population level to a sample is called sampling.

The first task in sampling is defining a target population of interest. At this step, the researcher needs to determine to which group he or she expects for the results of the study to apply. Decisions regard-

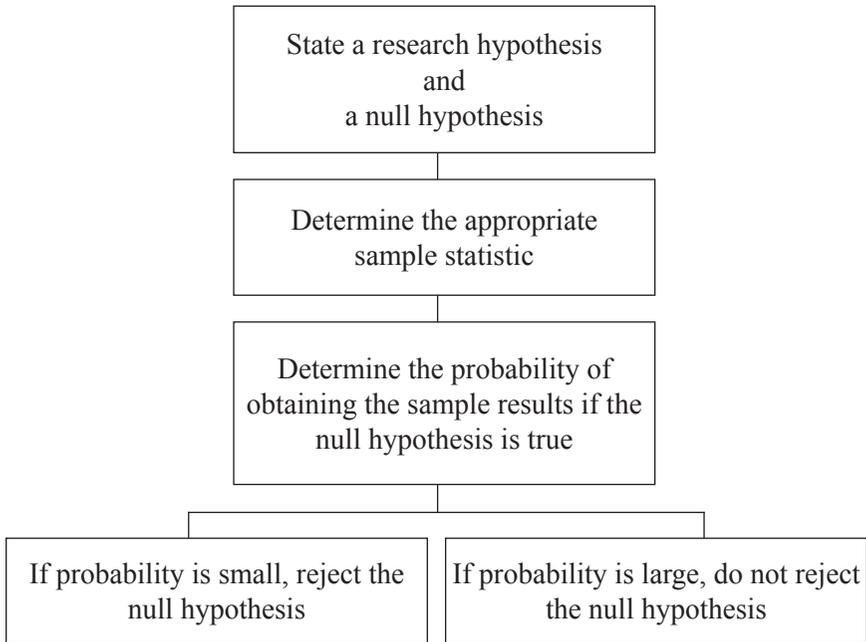


Figure 2. Summary of hypothesis testing logic. Adapted from *How to Design and Evaluate Research in Education* (8th ed.), by J. R. Fraenkel, N. E. Wallen, and H. H. Hyun, 2012, New York, NY: McGraw-Hill, pp. 228–229. Copyright 2012 by McGraw-Hill.

ing from which population to sample are typically derived from the research question. A population can be a large group of people, such as the residents of New York City, or a small and direct population, such as third grade students in Ms. Cha’s class in Gahanna Elementary School. The population would then represent all of the people who make up that specific description.

A researcher may use several strategies to select a sample from the targeted population. Sampling strategies are typically categorized into either random sampling or nonrandom sampling. Random sampling is when every member of a population has an equal chance of being selected into the sample (Fraenkel et al., 2012). The basic idea with random sampling is the group of individuals who are selected for the sample will have many of the same attributes or characteristics as normally distributed in the population from which it was drawn (Fraenkel et al., 2012). This is especially effective when the sample is large. On the other hand, nonrandom sampling is char-

acterized by all individuals in the population not having the same opportunity to be included in the sample. In nonrandom sampling, the researcher may establish additional criteria that he or she wants the participants to meet to be included in the sample. In this way, not all individuals in the population have the same chance (and some may have no chance) of being included in the sample. Recent document analyses and literature reviews of research in physical education for persons with disabilities indicate nonrandom sampling procedures were used in the majority of studies, categorized as purposive, convenience, or recruited samples (Haegele & Porretta, 2015a; Karkalatsi, Skordilis, Evaggelinou, Grammatopoulou, & Spanaki, 2012).

Researchers may use many random and nonrandom sampling strategies to obtain samples for research purposes. For this tutorial, we discuss briefly the features of six sampling strategies used commonly in quantitative research studies: (a) *simple random*, (b) *stratified random*, (c) *systematic*, (d) *cluster*, (e) *convenience* or *accidental*, and (f) *purposive* (Gay et al., 1996; Thomas et al., 2005).

Simple Random Sampling

Simple random sampling is a random sampling procedure in which each individual in the population has an equal chance of being included in the sample. Fraenkel et al. (2012) suggested simple random sampling may be the best method to obtain a representative sample of the population, especially for large samples. In this strategy, a table of random numbers is used to ensure every member has an equal and independent chance to be included (Fraenkel et al., 2012; Thomas et al., 2005). A table of random numbers can be a large list of numbers with no predetermined order or pattern. The table can be used in several ways to select participants. Fraenkel et al. (2012) suggested creating a series of random six digit numbers. From there, they would use the first few digits from each number to decide on the identified individuals within a population for the sample. For example, a researcher needs to choose 50 participants from a population of 500. The researcher would assign numbers to each participant (i.e., 1 to 500) and then go through the table of random numbers and use the first three digits of each number to choose the participants. If a number was 324053, participant 324 would be included in the study. If numbers were too high (i.e., 650), the researcher would skip to the next number.

Once again, the purpose of simple random sampling is to select a random sample that represents the larger population (Thomas et

al., 2005). This strategy has a few disadvantages. First, it is not an easy sampling strategy to use because each member of the population must be identified (Fraenkel et al., 2012). In large populations (e.g., inhabitants of states), this is not possible. Second, although school-based research is essential to physical education research, randomly selecting participants in schools is typically problematic. It is not likely school administrators will allow researchers to break apart classes in the name of scientific inquiry. Third, simple random sampling should not be used if researchers wish to be sure that certain demographic groups are included in the sample in the same proportion as they are in the population (Fraenkel et al., 2012). This can be an issue with small sample sizes. For this objective, researchers should use stratified random sampling or purposive sampling.

Stratified Random

Stratified random sampling is a sampling procedure in which the population is divided based on a chosen characteristic prior to sampling (Thomas et al., 2005). Participants who are included in a specified demography, or strata, are selected in the same proportion as they exist in the population, or the desired proportion for a study.

The following steps would be used in stratified random sampling. First, the researcher identifies the target population. Second, the researcher determines what characteristic he or she wants to stratify the sample based on and determine what percentage of that characteristic is present in the population. Next, the researcher creates a table of random numbers, which will include strata of each desired characteristic. Finally, the researcher uses the table of random numbers to select the sample, being sure to include a predetermined percentage of participants from each of the desired strata.

Stratified random sampling may be particularly useful in survey research, as researchers typically want to find a large representative sample (Thomas et al., 2005). For example, in conducting an online survey study, Beamer and Yun (2014) used a stratified national random sample of 3,000 public schools to examine the beliefs and self-reported behaviors of 233 general physical education teachers from 12 states across six geographic regions (two states per region sampled randomly) in the United States. The strata were states across regions. Advantages of stratified random sampling are it increases the likelihood of obtaining a representative sample and almost ensures important attributes of individuals are included in the same proportion as they naturally exist (Fraenkel et al., 2012). Perhaps the

biggest disadvantage of stratified random sampling is the amount of effort it takes the researcher to perform stratified random sampling correctly (Fraenkel et al., 2012).

Systematic

Systematic sampling is categorized as a nonrandom sampling method because all members of the population do not have an equal chance to be selected (Fraenkel et al., 2012; Gay et al., 2006). This method can be used when the population from which the researcher is sampling is too large and assigning a numeric identification number would be too time consuming (Thomas et al., 2005). Using systematic sampling methods, the researcher would select every n th (e.g., 12th, 7th, 122nd) individual in a list of potential participants. A method that is typically paired with systematic sampling is using a random start. A random start includes randomly selecting a starting point in the first few participants and then selecting every n th participant from there.

Several other terms are associated with systematic sampling. First, a sampling interval is the distance in a list between each of the participants selected for the sample (Fraenkel et al., 2012). If a researcher chooses a participant every 10th person, the sampling interval would be 10. Second, a sampling ratio is the proportion of the population that is included in the sample. If the population is 1,000 individuals and the researcher chooses 100 participants (one out of every 10th person), the ratio would be .10.

A benefit of systematic sampling is the selection process is simple (Gay et al., 2006). However, it is important for the researcher to inspect the list of potential participants carefully prior to sampling using a systematic sampling technique. If researchers are in an educational setting, it may not be uncommon to receive lists in order of grade point average (GPA), homeroom, or seat order in classes. Researchers should inspect the list for any pattern that could accidentally coincide with the sampling interval (Fraenkel et al., 2012). This type of bias is called periodicity. For example, Leites, Bastos, and Bastos (2013) used a systematic sampling procedure to obtain a representative sample ($n = 967$) of adolescents from South Brazil to estimate the prevalence of insufficient physical activity levels.

Cluster

Cluster random sampling is a technique researchers use when simple random and stratified random sampling strategies are not

available options. This can be when a list of all of the members of a population of interest is unavailable or if participants have predetermined groups, such as attempting to conduct research in schools. Occasionally, when it is not possible to sample individual participants in a school, it may be more likely to sample an entire intact class. The selection of classes, or clusters, instead of individual participants is called cluster random sampling (Fraenkel et al., 2012). For example, Hogan, McLellan, and Bauman (2000) randomly sampled one class from a collection of 115 target schools to determine the health promotion needs of students with self-reported disabilities in New South Wales.

Cluster random sampling is similar to simple random sampling except the researcher selects groups instead of individuals (Fraenkel et al., 2012). Benefits of cluster random sampling include it (a) is available to use when simple random sampling is impossible because of group contexts, (b) allows for research to be conducted in schools, and (c) can be less time consuming than simple random sampling (Fraenkel et al., 2012). However, with cluster random sampling, the chance is greater that the sample is not representative of the population.

Convenience or Accidental Sample

Researchers prefer random samples or systematic samples because of the high likelihood of representing the population accurately. However, it is not always possible to obtain a random or systematic sample of participants. In cases like this, researchers may recruit a convenience sample. *Convenience sampling* is selecting a group of individuals based on them being available for the study (Fraenkel et al., 2012). For example, Obrusnikova and Dillon (2011) used a convenience sampling method for their study. They contacted licensed physical education teachers through a list of e-mail addresses of individuals who had recently passed the Adapted Physical Education National Standards examination to complete survey questionnaires about challenges in teaching children with autism spectrum disorders (Obrusnikova & Dillon, 2011). This sampling strategy has the major disadvantage that there will likely be a bias in the sample. When using convenience sampling methods, researchers will likely choose participants who are either the closest in distance or the most accessible.

It is unlikely for convenience samples to be representative of the entire target population (Fraenkel et al., 2012). In this case, re-

searchers should be especially diligent in providing full descriptions of the participants in the study. As noted earlier, nonrandom sampling procedures, such as convenience or accidental sampling, are common in physical activity research, particularly pertaining to individuals with disabilities (Haegele, Lee, & Porretta, 2015; Haegele & Porretta, 2015a; Karkaletsis et al., 2012). This may be because of the difficulty in obtaining a sample of individuals with disabilities for studies, so researchers recruit individuals who are most available to them.

Purposive

Purposive sampling is another nonrandom sampling procedure in which researchers use their judgment to select a sample that they want. This is different than convenience sampling as researchers do not simply choose whoever is available for the study (Fraenkel et al., 2012). Rather, using purposive sampling, the researcher deliberately identifies criteria for selecting the sample (Gay et al., 2006). For example, in previous work, researchers (Haegele & Porretta, 2015b) have used a purposive sampling procedure for a study to seek validation information for audio pedometers. In this study, participants were selected based on residential status at a school for the blind, their ability to wear two pedometers simultaneously, and not having an ambulation-related disability (Haegele & Porretta, 2015b).

A primary strength of purposive sampling is the researchers can target attributes within a specific population and obtain a sample of individuals with those attributes. This strategy is especially beneficial when targeting populations that may be unique to the larger population because the researcher can pinpoint relevant attributes during sampling. However, purposive sampling has several weaknesses as well. Specifically, researchers may make errors in determining selection criteria of the sample or misestimate the representativeness of the sample that they select (Fraenkel et al., 2012; Gay et al., 2006). A further disadvantage of purposive sampling is that it limits the number and type of inferential statistics that are available to analyze the data.

Experimental Research Designs

The objective of experimental research is to establish cause-and-effect relationships (Thomas et al., 2005). To do so, the researcher manipulates the independent variable to judge its effect on the dependent variable (Fraenkel et al., 2012; Thomas et al., 2005). This

action allows experimental researchers to go beyond descriptive and correlational information and determine what causes the phenomenon to occur (Fraenkel et al., 2012). To establish a cause-and-effect relationship (a) the cause must precede the effect in time, (b) the cause and effect must be correlated with each other, and (c) the relationship between cause and effect cannot be explained by another variable (Thomas et al., 2005). Another characteristic of experimental designs is they typically involve at least two groups of participants. One group acts as the experimental group, who receives the intervention, and the other group acts as the comparison group (may receive an alternative intervention) or as the control group, who does not receive the intervention (Fraenkel et al., 2012).

In experimental research, one of the most important concepts is the control, or elimination or minimization, of threats to the validity of the results. Threats to validity can effect both *internal validity* (i.e., the degree to which observed differences on the dependent variable can be attributed directly to the independent variable) and *external validity* (i.e., the degree to which results are generalizable). To gain high internal validity, the researchers must control for all variables to eliminate other explanations for change (Thomas et al., 2005). When the researcher does this, though, the study will lose degrees of external validity because of the lack of ecological resemblance. Common threats to internal validity include (a) history, (b) maturation of participants, (c) testing effects, (d) instrumentation, (e) statistical regression, (f) selection bias, (g) experimental mortality, (h) selection–maturation interaction, and (i) expectancy (Thomas et al., 2005). One way to control for threats to internal validity is through randomization, which is discussed briefly in the next section on true experimental research. Placebos, blind, and double-blind studies are also strategies useful in controlling for threats to internal validity, but are not commonly used in physical education research. Some threats to internal validity, such as experimental mortality, are uncontrollable.

As with internal validity, external validity has several threats that can affect the ability for researchers to generalize results to other participants or settings. External validity threats include (a) reactive or interactive effects of testing, (b) interaction of selection bias and the experimental treatment, (c) reactive effects of experimental arrangements, and (d) multiple treatment interference (Thomas et al., 2005). External validity is typically controlled by selecting a sample

of participants who provide an equitable representation of the larger population (Thomas et al., 2005).

The way in which researchers choose to control variables and which variables they value in controlling affect the research designs they use. In experimental research, two broad categories are true experimental designs and quasi-experimental designs. In the following sections, we describe essential components of true experimental and quasi-experimental research designs.

True Experimental Research

Research designs are typically considered to be *true experimental* whenever they include randomly formed experimental and comparison or control groups, which allow the researcher to assume they were equivalent at the outset of the study (Thomas et al., 2005). Key aspects of random assignment include (a) it must occur prior to the experiment; (b) it must be a process of assigning individuals to groups, not an outcome of distribution; and (c) the groups that are formed are different only from chance (Fraenkel et al., 2012). The power of random assignment is that it controls for extraneous variables of which the researcher is or is not aware (Fraenkel et al., 2012). Random assignment controls for threats to validity including history, maturation, testing, statistical regression, selection bias, and selection–maturation interaction (Thomas et al., 2005).

Although true experimental research may be viewed as the most powerful style of group design research, it is not common in physical education research settings. One reason for this is researchers cannot randomize participants in an applied setting. In addition, the chance of providing a harmful treatment, or withholding a powerful treatment, causes ethical concerns in school-based settings. A third reason, pertaining more specifically to the education of individuals with disabilities, is the difficulty in acquiring a large enough sample of homogeneous participants with a specific diagnosis to conduct a study. These reasons lean more support for the use of quasi-experimental research (or single-subject designs) to evaluate treatments for students in schools.

Quasi-Experimental Research

The prefix *quasi* means “having some resemblance to a given thing” (Gove, 1971, p. 1861). Based on this definition, *quasi-experimental* research has some resemblance of true experimental research. According to Thomas et al. (2005), not all group design

research fits clearly into the category of true experimental design. In quasi-experimental designs, the design of a study is fit into the settings more likely to resemble real-world applications and as many threats to validity as possible are still being controlled (Thomas et al., 2005).

Because most real-world applications (e.g., physical education classes, schools in general) do not allow researchers to assign participants to groups randomly, randomization is usually the aspect of true experimental designs that is lost in quasi-experimental designs. When random assignment is not possible, researchers must rely on other techniques for controlling threats to validity (Gay et al., 2006). One such quasi-experimental design is the nonequivalent control group design. This design is similar to the experimental pretest–posttest design. However, rather than randomly assigning individuals to groups, researchers assign intact groups to different treatments (Gay et al., 2006). When assigning intact groups to different treatments, researchers should remember the appropriate unit of analysis is almost always the group, rather than the individual participants (Silverman & Solmon, 1998). Therefore, a larger number of groups may be needed to obtain sufficient power while analyzing data (Thomas et al., 2005). Other examples of quasi-experimental designs include counterbalanced, ex post facto, time series, and matching-only designs. The researcher may never control for internal validity as well in quasi-experimental designs as in true experimental designs, but quasi-experimental designs allow researchers to conduct investigations when true experimental designs are not feasible (Thomas et al., 2005). However, because the control and treatment groups may be different in unknowable ways, several alternative hypotheses may be stated to explain observed results in addition to experimental manipulation.

Summary and Implications

To conduct, analyze, and present research of high quality in the quantitative paradigm, researchers must have an understanding of basic assumptions. Therefore, it is essential for emerging professionals who have an interest in research pertaining to physical education for individuals with and without disabilities to be able to access this knowledge. In this tutorial, we have provided readers with basic information for understanding the quantitative research paradigm. This information has included the logic of hypothesis testing and sampling and key aspects of true and quasi-experimental research

designs commonly used in quantitative studies. In addition, we have provided relevant examples of the use of these components from the physical education and adapted physical education literature. This tutorial should help readers better understand basic concepts and principles of quantitative research methodology for the conduct of school-based research in physical education.

References

- Beamer, J. A., & Yun, J. (2014). Physical educators' beliefs and self-reported behaviors toward including students with autism spectrum disorder. *Adapted Physical Activity Quarterly*, *31*, 362–376. doi:10.1123/apaq.2014-0134
- Cervantes, C., & Porretta, D. (2013). Impact of after school programming on physical activity among adolescents with visual impairments. *Adapted Physical Activity Quarterly*, *30*, 127–146.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). New York, NY: McGraw-Hill.
- Gay, L. R., Mills, G. E., & Airasian, P. (2006). *Education research: Competencies for analysis and applications* (8th ed.). Upper Saddle River, NJ: Pearson Education.
- Gove, P. B. (1971). *Webster's third new international dictionary*. Springfield, MA: G. & C. Merriam.
- Haegle, J. A., Lee, J., & Porretta, D. L. (2015). Research trends in Adapted Physical Activity Quarterly from 2004 to 2013. *Adapted Physical Activity Quarterly*, *32*(3), 187–205. doi:10.1123/APAQ.2014-0211
- Haegle, J. A., & Porretta, D. L. (2015a). Physical activity and school-aged individuals with visual impairments: A literature review. *Adapted Physical Activity Quarterly*, *32*, 68–82. doi:10.1123/apaq.201-0110
- Haegle, J. A., & Porretta, D. L. (2015b). Validation of a talking pedometer for adolescents with visual impairments in free-living conditions. *Journal of Visual Impairment & Blindness*, *109*(3), 219–223.
- Hathaway, R. S. (1995). Assumptions underlying quantitative and qualitative research: Implications for institutional research. *Research in Higher Education*, *36*, 535–562.

- Hodge, S. R., Davis, R., Woodard, R., & Sherrill, C. (2002). Comparison of practicum types in changing preservice teachers' attitudes and perceived competence. *Adapted Physical Activity Quarterly*, *19*, 155–171.
- Hogan, A., McLellan, L., & Bauman, A. (2000). Health promotion needs of young people with disabilities: A population study. *Disability and Rehabilitation*, *22*, 352–357.
- Karkaletsis, F., Skordilis, E. K., Evaggelinou, C., Grammatopoulou, E., & Spanaki, E. (2012). Research trends in adapted physical activity on the base of APAO journal (2006–2010). *European Journal of Adapted Physical Activity*, *5*(2), 45–64.
- Leites, G. T., Bastos, G. A. N., & Bastos, J. P. (2013). Prevalence of insufficient physical activity in adolescents in South Brazil. *Brazilian Journal of Kinanthropometry and Human Performance*, *15*, 286–295. doi:10.5007/1980-0037.2013v15n3p286
- Levin, J. (1983). *Elementary statistics in social research* (3rd ed.). New York, NY: Harper & Row.
- Obrusnikova, I., & Dillon, S. R. (2011). Challenging situations when teaching children with autism spectrum disorder in general physical education. *Adapted Physical Activity Quarterly*, *28*, 113–131.
- Pringle, R. (2000). Physical education, positivism, and optimistic claims from achievement goal theorists. *Quest*, *52*(1), 18–31.
- Silverman, S., & Solmon, M. (1998). The unit of analysis in field research: Issues and approaches to design and data analysis. *Journal of Teaching in Physical Education*, *17*, 270–284.
- Thomas, J. R., Nelson, J. K., & Silverman, S. J. (2005). *Research methods in physical activity* (5th ed.). Champaign, IL: Human Kinetics.
- Trochim, W., & Donnelly, J. (2006). *The research methods knowledge base* (3rd ed.). Mason, OH: Atomic Dog.
- Vigo-Valentín, A., Bush, K., & Hodge, S. R. (2014). Daily physical activity behavior patterns of Hispanic adolescents in Puerto Rico. *Journal of Physical Activity & Health*, *11*, 1212–1218. doi:10.1123/jpah.2012-0169
- Wilkinson, M. (2013). Testing the null hypothesis: The forgotten legacy of Karl Popper? *Journal of Sports Sciences*, *31*, 919–920. doi:10.1080/02640414.2012.753636