The Pennsylvania State University

The Graduate School

College of Education

CLICKER STUDIES: REVIEW OF THE USE OF EDUCATIONAL PSYCHOLOGY PRINCIPLES, PROCEDURES, AND ITEMS

A Thesis in

Educational Psychology

by

Michelle L. Hepfer

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2018

The thesis of Michelle L. Hepfer was reviewed and approved* by the following:

Jonna M. Kulikowich
Professor of Education, Educational Psychology
Thesis Advisor

Bonnie J. Meyer
Professor of Education, Educational Psychology

Peggy N. Van Meter
Associate Professor of Education, Educational Psychology
Professor in Charge, Educational Psychology

*Signatures are on file in the Graduate School

ABSTRACT

Clickers have been used in classrooms as a tool to improve learning and affect for students. However, administrative procedures such as timed responses, peer discussion, and type of feedback carried out during studies on the effectiveness of clickers differ greatly because instructors differ in administrative preferences surrounding the use of clickers in their classrooms. In addition, many publications fail to provide descriptions or examples of the clicker items used in their studies. The differences in procedures and lack of information contribute to inconsistent and inconclusive results on the effectiveness of clickers for increasing student learning. This review assesses studies for their use of principles from the field of educational psychology to support student learning in conjunction with clickers; determines inconsistencies in the literature of the administrative procedures and types of feedback; and, analyzes descriptions and examples of clicker items asked. It was found that the most common principle from educational psychology applied in the studies was the use of multiple external representations (MERs) with the clicker questions. It was also found that the majority (n=104) of studies neither use nor address the use of peer discussion, re-voting, timed responses, or feedback when using clickers. Only 47% of the studies provided example items; of those, most possessed major design flaws based on empirically validated guidelines. Future recommendations include determining which administrative procedures consistently produce favorable results and a call for research on the effectiveness of using MERs with clicker questions.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

Introduction

Instructors of large lecture-style courses in post-secondary institutions have struggled with challenges concerning student engagement and learning outcomes (Addison, Wright, & Milner, 2009). Clickers, known by many names (e.g., personal response systems (PRS), instant response systems (IRS), audience response systems (ARS), classroom communication systems (CCS)), have grown popular in recent years for their proclaimed abilities to increase student academic achievement and improve student engagement in large, lecture-style courses (Mollborn & Hoekstra, 2010). Clickers are a classroom technology system that include computer software (e.g., TurningPoint, iClicker) and mobile devices used for formative assessment. Typically, an instructor will project a multiple-choice question on a screen, and the students will use a small, remote-like device called a "clicker", a mobile device such as a smartphone or iPad, or a computer to wirelessly submit an answer to the question. The instructor receives students' answers instantly, and the software aggregates the data into a bar chart. Next, the instructor can use this information to decide how to proceed with the lesson.

Despite the rise in popularity and initial proclamation of improving student learning, the literature concerning learning gains has produced mixed results. A recent meta-analysis (Hunsu, Adesope, & Bayly, 2016) found a statistically significant, small, positive effect size on cognitive outcomes across 86 studies. This meta-analysis also reported a statistically significant positive effect size for non-cognitive outcomes (e.g., students' perceptions and affect) across 25 studies.

These results indicate that using clickers, overall, can increase student-learning outcomes that relate to positive student affect toward courses.

Although the results of the recent meta-analysis favor the effectiveness of clickers, some questions remain about how studies using clickers are conducted. Methodological designs of studies implementing clickers into instruction have been problematic and inconsistent, making the effectiveness of clicker use unclear (Liu et al., 2017).

Studies also differ in administrative procedures. For example, some instructors allow students to discuss clicker questions with peers before submitting an answer (e.g., Such, Criado, & Garcia-Fornes, 2015; Willoughby & Gustafson, 2009) unlike studies where students were required to submit an answer without discussion (e.g., Aurus & Bix, 2007; Fallon & Forrest, 2011). Studies also differ in the amount of time students are allotted to answer a clicker question and whether the students are able to vote on the same question twice (re-voting). Finally, studies differ in the level of questions asked of students to answer using clickers. For example, Perez et al. (2010) focused on using higher-order questions, requiring students to integrate and apply their knowledge to answer the clicker questions. In contrast, Mollborn and Hoekstra (2010) reported using only low-level comprehension questions.

In addition to inconsistencies in administrative procedures, many studies have focused on affective outcomes over learning gains; reported only anecdotal evidence to support their claims; and, have failed to conduct studies free from internal threats to validity (Kay & LeSage, 2009). Recent literature reviews have not extensively or critically examined the methodologies and administrative procedures of the studies reviewed (e.g., Hunsu et al., 2015; Kay & LeSage, 2009; Liu et al., 2017). Consequently, these reviews have compared the results of studies with

incompatible research designs, administrative procedures, outcome measures, domains, and uses of the clicker technologies.

In relation to differences in administrative procedures, studies have been inconsistent in reporting the types of formative assessment questions presented to students when using clickers and providing examples of clicker questions used. There have also been concerns about whether or not the "traditional" clicker question formats (i.e., multiple-choice and true/false) are able to activate and assess higher-order thinking skills (Liu et al., 2017). As this review will further illustrate, most studies do not report question format, question design, or provide example items.

In light of the numerous inconsistencies in the literature, this review will seek to report:

1. administrative procedures during clicker use.
2. implementation of feedback to students.
3. assessment of clicker items for design and level of thinking required.
4. use of educational psychology learning theory principles in clicker assessment items.

Clickers and Implementation

Clickers in classrooms take on a wide-range of hardware devices. Hand-held, remote-like devices using radio frequencies or infrared to communicate information from students to a receiver base connected to the instructor's computer are still common today (e.g., iClicker), but many institutions have transitioned to web-based questioning systems (i.e., using Wi-Fi and mobile devices; Turning Technologies, 2017). Traditionally, clickers have been used with primarily multiple-choice questions as a result of the limitations of the devices and software. With emerging technology (e.g., QT response device by Turning Technologies, and iClicker Reef mobile application), open-ended and other complex questioning methods, such as matching, can be used (Macmillan Learning, 2017; Turning Technologies, 2017).

In a typical lecture-style classroom employing clicker technology, an instructor will project a question on a screen that is often integrated into a PowerPoint® presentation, or on students' devices (e.g., iPhones, tablets, laptops). The questions can be used for a variety of reasons: 1) to assess students' prior knowledge; 2) to stimulate thinking; 3) to assess students' learning of new content (i.e., formative assessment); or, 4) for administration of quizzes and tests (i.e., summative assessment). The instructor can decide whether or not to use time constraints on student responses, allow or disallow peer discussion before answering, provide feedback about the responses, and decide the manner in which feedback is provided. Students use their devices to wirelessly and anonymously submit their answers to the question, which is often called "voting". Often, instructors will display the software's aggregated voting data, typically in the form of a bar chart, to the students and provide the correct answer to the question (e.g., Blasco-Arcas, Buil, Hernández-Ortega, & Sese, 2013; Chien, Lee, Li, & Chang, 2015; Desrochers & Shelnutt, 2012; Gauci, Dantas, Williams, & Kemm, 2009). Then, instructors can have the students engage in discussion about the answer or provide an explanation. Instructors may also decide to allow students to re-vote on the same question following discussion. Finally, the instructor decides how to proceed with the lesson given the voting results.

Many studies have examined clickers' effectiveness as related to higher learning gains (cognitive outcomes) as well as affective effects on students (e.g., motivation, engagement, participation, attendance). Gauci et al. (2009) collected survey data from students using clickers in undergraduate physiology courses. They found that students felt engaged (83%), intellectually stimulated (85%), and motivated to think (89%) as a result of using the clickers. Similarly, Friedline, Mann, and Lieberman (2013) found that students in an undergraduate social work course felt more comfortable answering formative assessment questions using clickers than using

hand raising. The students also felt that using clickers was enjoyable, easy, engaging, and helpful for focusing. The authors also concluded that using the clickers in the course helped to improve class discussion by facilitating greater participation. Numerous additional studies have also reported positive non-cognitive outcomes as a result of using clickers based on student survey data (e.g., Auras & Bix, 2007; Chen & Lan, 2013; Cheung, Chan, Wan, & Ng, 2015; Clauson, Alkhateeb, & Singh-Franco, 2012; Grzeskowiak, Thomas, To, Reeve, & Phillips, 2015).

While studies of affective outcomes have been overwhelmingly positive, results regarding cognitive outcomes have been mixed. Gauci et al. (2009) found statistically significant improvement on final-exam performance, with the clicker group scoring an average of 72% correct and the non-clicker group scoring an average of 66% correct. Mayer et al. (2009) also found greater learning gains for students using clickers in an undergraduate educational psychology course than students who answered questions on paper or a control group that did not answer any questions. Students who used clickers to answer adjunct questions during the class lecture performed significantly better (one-third grade-point average) on the final exam than students who answered questions on paper (non-clicker) and students who did not answer any questions in class (control). The non-clicker and control groups' scores on the final exam did not differ significantly from each other.

In contrast, Fortner-Wood, Armistead, Marchand, and Morris (2013) found no significant differences in final course grades of a lifespan development class, where the clicker group had an average of 75% at the end of the semester and the comparison group (traditional lecture) had an average of 77% (SD = .10 and .11, respectively). Anthis (2011) found that students who used clickers in one section of a lifespan development course performed slightly worse on a class

exam than students who did not use clickers (M = .78, and M = .82, respectively), though this difference was not significant. A second exam in the same course revealed that the clicker group performed only slightly better than the non-clicker group (both M=.75), though this difference also was not significant. As stated above, discrepancies in the results of these studies may be due to lack of consistency in research methodology and classroom implementation, as well as differences in the nature of the domains in which clickers are researched (Hunsu et al., 2016).

Formative Assessment and Feedback

Formative assessment is essential for collecting evidence of student mastery, promoting learning, and informing teaching decisions (Stiggins, 2005). A comprehensive literature review conducted by Black and Wiliam (1998) analyzed the effects of formative assessment on student learning. The review found that students who are frequently assessed show higher learning gains than those students who are not. The authors reported that formative assessment questions help students to focus their attention and emphasize the importance of the concepts being learned. In addition, no negative effects on learning were encountered after enhancement of formative assessment practices. However, some weaknesses were observed in the implementation of formative assessment practices. For example, many instructors tend to use only low-level knowledge and comprehension questions, fail to provide feedback or opportunities for peer discussion, and over-emphasize grades over learning.

In addition to implementation of formative assessment practices, instructors should consider the manner in which they provide feedback to the students. Feedback, which is information about a learner's performance that is intended to improve thinking or ability (Shute, 2008), is important for the correction of errors and is beneficial to future student performance (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). Good feedback can be a significant factor for

improving and motivating student learning. A meta-analysis by Bangert-Drowns et al. (1991) and a literature review by Shute (2008) highlighted learning gains as a result of feedback. On average, feedback had a small, positive effect on achievement with an increase of approximately one-fourth of a standard deviation. Both reviews indicated that feedback is best when it is immediate, frequent, and corrective. In addition, Butler, Godbole, and Marsh (2013) found that explanation feedback, feedback in which the instructor explains why a particular answer is correct, improves student performance over providing the correct answer alone.

Given these important findings of formative assessment and feedback, it is essential that instructors using clicker technology are conscious of how they are implementing clickers as an assessment tool and providing feedback. One critique by Black and Wiliam (1998) explained that instructors tend to use low-level questions, those which assess declarative knowledge, on formative assessment measures. Instead, instructors should supply high-level thinking questions that address conceptual knowledge and require cognitive tasks such as applying and analyzing as much as possible. In addition, instructors should be mindful that their feedback is immediate, frequent, and focused on correcting student errors. Instructors who wish for students to gain deeper understanding should also consider implementing explanation feedback, or explain why a particular answer is correct, after answering clicker questions (Butler et al., 2013).

Educational Psychology

Clickers can be a useful tool in the classroom, but to ensure the highest learning gains will be reported instructors should consider implementing basic instructional principles researched and validated by educational psychologists. A few principles can easily be applied to the formative assessment environment that clickers introduce: activating prior knowledge, using

multiple external representations (MERs), addressing misconceptions, and promoting self-regulated learning.

Effects on student learning after activating prior knowledge have been studied for nearly a century by educational psychologists. Prior knowledge has the ability to influence students' interpretations of new information and guide their predictions. It is also a valuable resource that can be built upon to facilitate deeper learning (Kostons & van der Werf, 2015; Mayer, 1979). Instructors can present clicker questions aimed at targeting applicable prior knowledge at the beginning of a new lesson. By bringing relevant information to students' working memory, students have the opportunity to make connections between the old and new information, which results in deeper understanding and increased retention (Kostons & van der Werf, 2015).

Use of MERs in instruction has also been studied extensively. Based on Cognitive Load Theory (Sweller, van Merrienboer, & Paas, 1998) and the Cognitive Theory of Multimedia Learning (Mayer, 1997), MERs are able to overcome the capacity limitation of working memory by incorporating multi-modal sources of information (e.g., verbal/visual, iconic/symbolic; Ainsworth, 2006). MERs can play a major role in the acquisition of learning as they often require *abstraction*, or creation of a mental model that reflects the underlying structure of the representation; *extension*, or application of knowledge from one representation to another; and *relational understanding*, or associating representations without reorganizing or extending knowledge. All such processes facilitate deeper understanding of complex information (Ainsworth, 2006). Instructors can integrate MERs into clicker questions by writing questions that require students to analyze an external representation such as a photograph, diagram, chart, or graph in order to answer the item correctly.

Another related consideration is the domain in which the clickers are used. That is, instructors should consider best assessment practices for their specific domain (e.g., Alexander, Murphy, & Kulikowich, 1998). For example, clicker questions that require students to analyze visual aids such as photographs and diagrams in an earth science class is an important consideration for the domain as earth science involves activities such as classifying rock formations, reading maps, and identifying types of clouds. Another example is consideration for the domain of English as a second language. The English language contains an abundance of synonyms which may be confusing for English language learners. Therefore, designing questions that address the intricacies and nuances surrounding different synonyms is an important consideration for item writing in an English language class.

Multiple-Choice Question Design Practices

Multiple-choice questions can assess many high-level thinking abilities including verbal, numerical, mechanical, and abstract reasoning, spatial relations, clerical speed and accuracy (Ghiselli, Campbell, & Zedeck, 1981), as well as evaluating, predicting, and problem solving (Haladyna, 1997; Haladyna & Downing, 1989a). Questions designed to assess high-level thinking can help students learn not only through retrieval practice, but will also require students to make new connections within the information they have already learned, that helps to strengthen their understanding and promoting problem solving. Using the Taxonomy of Cognitive Objectives (i.e., Bloom's Taxonomy), one can classify assessment items to show the level of thinking required by students to answer the question or solve the problem (Airasian & Miranda, 2002). Typically, objectives and items that are classified in the categories from *Understand* to *Create* (i.e., understand, apply, analyze, evaluate, and create) are considered to be

the most important (Krathwohl, 2002). As such, instructors should focus on writing high-level (i.e., apply, analyze, evaluate, and create) assessment items to foster these skills.

In addition to consideration for the type of knowledge addressed by questions, guidelines for writing clear and effective multiple-choice questions have been in use since the 1960s (Adkins, 1960). Careful attention should be given when constructing the question itself (stem), the correct answer choice (key), and the incorrect answer choices (distractors). According to Ghiselli et al. (1981), many of these formatting recommendations are "a matter of common sense" (p. 427). Items should:

1. deal with one central thought.
2. use clear and precise language.
3. be brief, or economical in language.
4. avoid awkward wording or dangling phrases.
5. avoid irrelevant information.
6. include as much information in the stem as possible and keep answer choices brief.
7. be stated positively and avoid double negatives.
8. use plausible distractors.
9. have answer choices that are consistent in length, grammar, and structure.
10. avoid synonymous distractors.
11. randomize the position of the key among the distractors.

In addition to Ghiselli et al.'s principles, Haladyna and Downing (1989b) add: (a) use *quality* distractors, those that increase discrimination; (b) avoid using "none of the above," "all of the above," and "I don't know"; and (c) avoid complex, or Type K format. Complex, sometimes called Type K, format multiple-choice questions contain answer choices such as "both choices A

and B are correct." Type K questions may also provide a list of statements and answer choices such as "statements I and II are correct."

Haladyna and Downing (1989a) examined empirical validation of the above guidelines. They reviewed studies that examined item-level differences between well and poorly written items. In general, items that do not adhere to the design guidelines can unintentionally increase the difficulty of an item, decrease the discriminatory property of the item (i.e., the item's ability to distinguish between high- and low-achieving students), and decrease the reliability of the scores of the measure containing that item.

Varying the degree of correctness of the distractors is also an important consideration since multiple-choice items require that students must distinguish between answer choices that differ in relative correctness (Popham, 2011). For a four-option multiple-choice item, this means that the key is the *best* choice for completing the stem, or the *most* correct, one choice is close in correctness to the key, a third choice is farther away in correctness, and the fourth choice is the most incorrect among them. Varying the degree of correctness allows an instructor to assess the degree to which students understand the material, as they will need to distinguish between key features within each answer choice to determine the one that best completes or answers the stem (Gierl, Bulut, Guo, & Zhang, 2017).

Although the development of the formatting principles focused on summative assessments, the practices are still valid for the creation of reliable, fair, and quantifiable clicker questions (Beatty, Gerace, Leonard, & Dufresne, 2006; Ghiselli et al., 1981). Summative measures assess the results of instruction and are commonly used to make decisions about students' achievement (e.g., course grade, ability level, receiving a driver's license). Such measures often have high stakes for the student. In comparison, formative assessments,

especially those using clickers, may seem insignificant as they do not typically have high-stakes associated with their results. However, consider that clicker questions are often used by instructors to decide whether or not students understand the course material. As such, these evaluations may affect an instructor's decision to provide further explanation, instruction, or reteach the content. Since flawed questions have been shown to relate to item difficulty, item discrimination, and reliability of scores, the instructor can be gaining an inaccurate portrayal of student understanding if the clicker questions used possess design flaws. These implications can greatly impact a student's level of learning within a subject which can potentially impact his or her later stages of education, licensure exam performance, and career performance. Therefore, much care should be given to the quality of the clicker questions presented.

The following section describes the methods used for assessing the literature of 104 studies using clickers. Each study was analyzed for descriptions and examples of items, administrative procedures and implementation practices for using clickers, and the principles of educational psychology evoked to guide the use of clickers.

Chapter 2

Methods

This review analyzes the same studies included in the recent review conducted by Lui et al. (2017). Lui et al. located 124 articles by searching the Web of Science for educational research studies using the key terms "clickers," "instant response systems,", and "IRS". All articles were published between 2007 and 2015. This collection of articles was deemed appropriate for the present review because the articles were:

1. recent, to limit the number of articles and to assess reports using the most current technology and practices.
2. empirical, to focus only on studies that used an experimental or quasi-experimental design.
3. exhaustive, to find all relevant articles.
4. focused on the use of clickers in classrooms.
5. written in English.

After an initial overview, 20 studies were dropped from this review due to implementation issues, such as using clickers to play a game or evaluate peers, or methodological issues such as descriptive or conceptual publications. Ultimately, 104 studies were included in this review.

The contents of each publication were analyzed for administrative procedures (e.g., peer discussion, timed responses, re-voting, feedback), descriptions of clicker items, examples of items, use of educational psychology principles for item writing, and type of feedback provided

to the students (e.g., stating the correct answer, showing voting results, providing an explanation). Hypotheses, type of clicker software and devices used, results, and conclusions of the studies were also noted.

The author was assisted by four raters to code the studies. Of the 104 articles used in this review, 25 (24%) were coded by two raters (the author and a second rater). The raters used the articles' methods sections to code for administrative procedures and type feedback. The raters coded descriptions and examples of the items verbatim and descriptively for level of thinking required (e.g., low-level, apply, analyze) and the extent to which the questions follow best design practices (Ghiselli et al., 1981; Haladyna & Downing, 1989a, 1989b). The raters also analyzed the introduction and methods sections, example items, and references to determine the use of educational psychology principles for assessment. The raters searched the references for citations from high-impact journals in educational psychology, assessment, and related fields (e.g., *Journal of Educational Psychology*, *Educational Psychologist*, *Educational and Psychological Measurement*; a complete list can be found in the Appendix), and determined the purpose of the reference by searching for the citation in the article. The author of this review was the only rater for the remaining 79 (76%) articles.

Interrater agreement was established for the 25 articles that were coded by two raters. Initial interrater agreement reached 80.0% (70 errors made in 350 total codes). Many discrepancies occurred when information was overlooked by one of the raters, or the second rater misunderstood the information desired. No systematic errors were detected.

Chapter 3

Results

Administrative Procedures

Questions per class and positioning. The number of questions asked per class, lecture, or session were reported in 56 (54%) studies. Of those, the majority (82%) reported posing 1 to 15 questions per class. Other studies in which students took quizzes using clickers asked 15 to 50 questions per class. Two studies that took place in a computer lab using a simulated clicker environment used 64 questions. It is important to note that the studies varied in positioning of the questions throughout the lecture, though it was not explicitly coded. Some authors used clicker questions at the beginning of class to assess or activate students' prior knowledge, test them on the assigned reading, or review from the previous class (e.g., Morling, McAuliffe, Cohen, & DiLorenzo, 2008; Prunuske, Batzli, Howell, & Miller, 2012), while others asked all of the clicker questions at the end of class (Anderson, Healy, Kole, & Bourne, 2011; 2013). Most studies, however, dispersed the clicker questions throughout the lectures (e.g., Kim et al., 2015; McLaughlin, Gharkholonarehe, Khanova, Deyo, & Rodgers, 2015; Stowell & Nelson, 2007). Not all studies indicated when clicker questions were presented.

Peer discussion. Table 1 summarizes the major trends in administrative procedures. Forty studies reported that peer-to-peer discussion was always or sometimes allowed prior to answering the clicker questions. Of those, 14 only allowed discussion before re-voting, 16 always allowed peer discussion, and 2 studies manipulated peer discussion as a variable in the experimental conditions. The remaining studies allowed students to discuss if they desired or did

not indicate the circumstances under which discussion was allowed. Peer discussion that took place exclusively before re-voting was typically described by the authors as "convince your neighbor" discussion; students would enter a discussion with an answer in mind and would try to persuade their peer that they were correct (Brady, Seli, & Rosenthal, 2013a; Cardoso, 2011).

Table 1
Administrative Procedures

	Peer Discussion		Re-voting		Timed Responses	
	N	%	N	%	N	%
Yes	26	25.0	11	10.6	18	17.3
No	64	61.5	80	76.9	83	79.8
Sometimes	14	13.5	13	12.5	3	2.9

Re-voting. Re-voting occurs when the same clicker question is asked and voted on more than one time. The methods of the majority of the studies (see Table 1) did not give the students an opportunity to re-vote on a clicker question or did not indicate using this practice. The remaining studies used re-voting in various ways and for various reasons. Some studies had students re-vote on every question, every time, while a similar amount only used re-voting under certain conditions. An unsatisfactory distribution of students answering correctly was the most common reason why studies would occasionally use re-voting. The criterion for most studies was if fewer than 70-75% of students voted for the correct answer the first time. Other reasons for sometimes using re-voting were vaguely indicated as "as needed," or "after discussion" by the authors.

An example of a study in which students re-voted on all of the clicker questions was conducted by Morice, Michinov, Delaval, Sideridou, and Ferrieres (2015), who examined the effects of peer discussion between initial voting and re-voting. Participants were assigned to either a peer instruction (experimental) or individual (comparison) condition and used clickers to

answer 14 questions about chromatography. During the first time the question was presented, participants in both conditions answered individually. The initial voting results were presented, but no further explanation or instruction took place. The same question was presented again, and the participants in the peer instruction condition discussed the question in groups of four, while participants in the individual condition thought only to themselves before re-voting. The process was the same for all 14 clicker questions. They found that students in the peer instruction group outperformed those in the individual group on the 14 clicker items. Students in the peer instruction group also reported higher rates of participation and satisfaction with the class.

Timed responses. Twenty-one studies reported at least sometimes giving the students a restricted amount of time to answer the clicker questions; however, only 16 (76%) of those indicated the amount of time allotted. The most common amount of time given to answer a clicker question was about one minute, with some variations (e.g., 1-2 minutes, 30-90 seconds). The longest amount of time reported was five minutes and the shortest was two seconds.

Feedback. Since quality feedback has been shown to increase learning outcomes (Bangert-Drowns et al., 1991), it was determined to be a crucial piece in the administrative procedures of the clicker studies. Sixty-one studies indicated implementing some type of feedback. From simply stating the correct answer to full-class discussions, the types of feedback given varied greatly. Table 2 highlights the major themes observed. It is noted that many studies used more than one method of feedback. For example, one instructor first displayed the distribution of the voting results, provided the correct answer, encouraged peer discussion, then retaught if necessary (Brady, Seli, Rosenthal, 2013a). In this study, therefore, all four of those feedback types were coded.

Table 2
Feedback

Provided	N	%
Yes	61	58.7
No	4	3.8
Not Addressed	39	37.5

Most clicker software programs allow for the immediate display of aggregated voting data in a histogram or bar chart to the class. Of the studies reviewed, a majority of those that provided feedback displayed the voting results to the class. The second most common type of feedback was discussion, either with peers only or the whole class. Only 31.1% of the studies that addressed the use of feedback indicated that the instructor provided the correct answer to the students. Furthermore, two studies (Levesque, 2011; Lin, Liu, & Chu, 2011) indicated that the researchers explicitly did *not* provide students with the correct answer to the questions. Explanations from the instructors were the fourth most common type of feedback, and immediate feedback was fifth with ten studies reporting implementing feedback immediately after voting took place (see Table 3).

Since literature reviews on feedback indicate that feedback is best when it is immediate, explanatory, and corrective, studies that gave feedback immediately, provided an explanation, and indicated the correct answer, were likely to be using the best practices for feedback implementation. No study in the collection of articles analyzed in this review reported providing the correct answer, providing an explanation, and giving immediate feedback in conjunction.

Table 3
Type of Feedback

	N=61	%
Shown Voting Results	35	57.4
Discussion	29	47.5
Correct Answer	19	31.1
Explanation	17	27.9
Immediate	10	16.4
Reteaching	3	4.9
Individual	1	1.6

Note. Percentages add to over 100% because several articles used more than one method of feedback; percent reflects percentage of the number of articles that provided feedback.

Items

Format and descriptions. The majority of studies used multiple-choice format for their clicker questions (see Table 4). Typically, multiple-choice questions were posed with 4 answer choices, but some researchers used five, or even up to ten (e.g., Ludvigsen, Krumsvik, & Furnes, 2015). Although most studies used only multiple-choice questions, many studies used combinations of item formats, including true and false, numerical short answer, and fill-in-the-blank. Two studies also included Likert-type questions in order to collect data on student attitudes or opinions (Pagano & Paucar-Caceres, 2013; Stowell, Oldham, & Bennett, 2010). Thirty-five studies did not address the format of the clicker questions used.

Seventy-one (68%) studies provided a description of the clicker questions used. Descriptions varied from short and vague such as, "probed students' understandings of the ideas just presented" (Gök, 2011b, p. 55), to detailed and elaborate, "the clicker questions…focused on practical scenarios…and how to interpret a graph…. When possible, incorrect answer choices were written based on incorrect ideas stated…during previous interactions." (Smith, Annis, Kaplan, & Drummond, 2012, p. 2). More detailed descriptions often included information about

item format(s), level(s), sources (e.g., textbook item bank, validated instruments), knowledge type(s) measured, and distractor development.

Table 4
Item Formats

	N=104	%
Not Addressed	35	33.6
Multiple-Choice	65	62.5
Short Answer	7	6.7
True/False	6	5.8
Fill-in-the-Blank	2	1.9
Likert-Type	2	1.9
Ranking	1	1.0

Note. Percentages add to over 100% because several studies used more than one type of item format

Examples. In order to assess the specific items used in clicker studies, example items needed to be pulled from the articles. Forty-nine (47%) studies provided examples of the clicker questions in the publication or as supplemental material. Items were analyzed for level, design practices, and application of educational psychology principles. Table 7 summarizes the characteristics of the example items. It is important to note that not all of the example items could be assessed for level of thinking due to lack of context and/or expertise of the raters. Therefore, many of the levels reported relied on the author's description of the item. The most notable results include majority low-level questions and the most-used principle of multiple external representations. Example items were coded for using multiple representations when students needed to interpret a graph, map, chart, or photograph to answer the question. The most common design flaws in the examples included using answer choices "none of the above," "all of the above," or "I don't know," distractors with incompatible grammatical structure, implausible distractors, and varying distractor length (see Table 5).

Table 5
Characteristics of Example Items

Level	N=49	%
Low (e.g., know, understand)	22	44.9
Middle (e.g., apply)	13	26.5
High (e.g., analyze, evaluate, create)	5	10.2
Design Flaws		
"None of the above", "all of the above", "I don't know"	13	26.5
Incompatible Grammatical Structure	13	26.5
Implausible Distractors	12	24.5
Distractor Length	11	22.4
Complex Multiple-Choice	8	16.3
Too Long	5	10.2
Negatively Stated	5	10.2
Two or More Central Ideas	4	8.2
Unclear or Awkward	4	8.2
Synonymous Distractors	1	2.0

Note. Percentages add to over 100% because many articles provided more than one example with different levels, design flaws, or principles; some items could not be assessed for level due to lack of context or domain expertise; design flaws are based on multiple-choice examples only.

Educational Psychology Principles

The most common principle from educational psychology applied by researchers was the use of multiple external representations (MERS; e.g., charts, graphs, photos, and diagrams) within the clicker questions. For example, Prunuske, Batzli, Howell, and Miller (2012) provided an example item which used diagrams of gametes and chromosomes that students needed to analyze in order to answer a question about meiosis. Table 6 reports the principles that were detected in the example clicker questions provided in the publications or in supplemental material.

Table 6
Principles in Example Items

	N=49	%
MERs	14	28.6
Reasoning	3	6.1
Domain-Specificity	2	4.1
Activate Prior Knowledge	1	2.0

Note. Percentages add to less than 100% because not all example items used a principle from educational psychology.

In total, 15 studies reported using MERs in their clicker questions. Table 7 reports the principles that were addressed or described by the authors for use in the clicker questions. Percentages in this table are out of all 104 studies analyzed. Studies were included in this table only if the author described using MERs or explicitly stated that clicker questions contained pictures, graphs, diagrams; actual examples provided by the authors were not included in Table 7.

Table 7
Educational Psychology Principles Described by Authors

	N=104	%
MERs	15	14.4
Activate Prior Knowledge	6	5.8
Misconceptions	6	5.8
Domain-Specificity	2	1.9
Self-Regulated Learning	1	1.0
Questioning Strategies	1	1.0

Note. Percentages add to less than 100% because not all studies reported or described the use of educational psychology principles.

Six studies described using the clicker questions to activate students' prior knowledge that typically happened at the beginning of the class or lecture. Instructors often reported using information gathered from these types of questions to alter the lesson to fit the needs of the

students. In addition, many studies indicated that the students benefitted from these practices because the questions could "draw students' attention to their own limitations," (Gauci et al., 2009, p. 61). In other words, using clickers to activate prior knowledge facilitated metacognition as it became a means for students to assess their preexisting understanding of the concepts. Aside from the metacognitive benefits of activating prior knowledge, only one study addressed the cognitive benefits of activating prior knowledge on future learning. Goldberg, Price, Robinson, Boyd-Harlow, and McKean (2012) built an intervention design principle based on prior knowledge's ability to influence students' interpretations and predictions of new content, indicating that activating prior knowledge will also help facilitate students' understanding of the new content.

Two studies considered domain-specific characteristics of the subject in which the learning took place when writing the clicker questions. To illustrate, Lee and Feldman (2015) considered the domain of astronomy as it relates to the sun and moon by constructing questions that contained photographs of the observable phenomena about which the students were learning. For example, students were given a photograph of the sun on the horizon and told that it was taken near the north pole in December. The students were asked to indicate what time of day the photograph was taken. Cardoso (2011) also considered specific characteristics of the English language when constructing the clicker questions for English language learners in Brazil. Specifically, Cardoso considered that English contains many words that have the same or similar meaning, and that some are more rarely used than others. The questions posed to students in this study related directly to their English readings and were designed to test uncommon vocabulary by creating synonymous distractors of which students needed to identify the correct word used in their text.

References. To assess the field of educational psychology's impact on studies using clickers, each study's reference section was analyzed for references from top-cited journals in educational psychology (e.g., *Journal of Educational Psychology*, *Contemporary Educational Psychology*; see Appendix for full list), or related fields (e.g., *Cognitive Psychology*, *Journal of the Learning Sciences*). Forty-three (41%) studies contained at least one reference to an article in a top educational psychology journal. Each reference was analyzed for purpose. The results are reported in Table 8. Support for the theory behind the study, providing examples, and support for administrative procedures were the most common uses of the educational psychology journal references. For example, Yu, Chen, Kong, Sun, and Zheng (2014) had six references to four different journals that reference educational psychology principles: *Learning and Instruction* (2), *Educational Psychology Review* (2), *Educational Psychologist* (1), and *Journal of Educational Psychology* (1). All six of these references were used in the study's literature review under the cognitive load subsection. The references were found at the end of sentences in which the authors defined and explained cognitive load and established a rationale for addressing its implications in their study and were therefore coded as "theory."

Table 8
Use of Educational Psychology Journal References

Purpose	N=43	%
Theory	20	46.5
Example	14	23.3
Administrative	6	14.0
Testing	4	9.3
Instrument	3	7.0
Feedback	1	2.3
Item Writing	1	2.3
Other	6	14.0

Note. Percentages add to over 100% because some articles had multiple references serving different purposes.

The second most common use of references from top-cited journals was as an example of a study that also assessed the effects of clickers. The most cited study was by Mayer et al. (2009) from *Contemporary Educational Psychology*. This study was cited 18 times in the reviewed collection of studies for its findings of positive effects on learning outcomes, comparison of different questioning strategies, and application of the Generative Theory of Learning (e.g., Wittrock, 1990).

References which provided support for the administrative procedures used in the study were coded as "administrative." For example, Levesque (2011) used two references from the *Journal of the Learning Sciences* that supported the use of peer discussion to facilitate learning. As a result, Levesque allowed students to engage in peer discussion before answering clicker questions while participating in the study. She found a significant, positive correlation between clicker participation and exam performance.

References were coded as being used for "testing" if the reference informed the use of formative assessment. For example, Cardoso (2011) cited an article from the *Journal of Educational Psychology* in order to explain that observed learning gains may be linked to using adjacent questions and feedback, a basic use of clickers. Studies that used measurement tools (e.g., surveys or cognitive exams) from educational psychology references were coded as "instrument." If a reference from a top educational psychology journal was used to support the use or method of feedback, then "feedback" was coded.

Since one of the purposes of this review is to examine how clicker studies have applied educational psychology principles to the questions they pose to students, it was important to note if the reference from educational psychology directly influenced the writing of the clicker items used in the study. Only one study used a reference in such a manner. Lee and Feldman (2015)

referenced an article from *Cognitive Psychology* that found that learning an abstract concept may be facilitated by concrete representations. Therefore, the researchers created clicker items that included photographs that served as concrete representations. They used a survey to assess students' affect toward the class and found that the students believed that the photographs helped their learning, helped them to be more observant, and were more engaging.

Chapter 4

Discussion

In-class clicker questions are a formative assessment tool often used as an intervention to improve cognitive outcomes for students. However, many published studies on the effectiveness of clickers to produce positive cognitive or non-cognitive outcomes do not specify the administrative procedures used with the clickers (i.e., peer discussion, timed questions, re-voting), type of feedback given to the students about their answers, guidelines for constructing the clicker questions, or descriptions and examples of the clicker questions asked. Based on studies that provided examples of the clicker questions, it was found that instructors typically presented questions with flawed design. In addition, it was found that fewer than half of the studies on clickers use empirically validated instructional methods and influences from educational psychology and related fields.

Administrative Procedures

This analysis of studies on clicker effectiveness showed that studies differed greatly in the administrative procedures used to implement clickers. Many studies used peer discussion, re-voting, time limits, and feedback, or various combinations of these procedures. Variation in the ways researchers implemented these procedures were also detected across studies. In addition, many studies were not transparent about administrative procedures or did not address them in the publications.

Studies in this literature review included field studies and laboratory-type studies. Some field studies were conducted in semester-long courses, where students used clickers every day,

while others only used clickers for one or two weeks. Other field studies used clickers for only one presentation. Laboratory studies used simulated class environments, and therefore were not authentic to how students and instructors typically use clickers in classrooms.

Considering these procedural differences and variations, the lack of consensus on the effectiveness of clickers is understandable. Past literature reviews and meta-analyses have not distinguished between these differences in implementation of clickers. Therefore, it is not surprising that the results have been inconclusive or in favor of clickers by only a small margin.

Items

Although most studies used multiple-choice items, some used other formats or a combination of item formats. Authors also described their items in different ways, claiming that they addressed different types and levels of knowledge or assessed understanding gained during different instructional activities (e.g., reading before class, prior knowledge, lecture content). Questions were also presented at varying times during instruction across studies.

In addition, flaws in assessment design were common among the example items. Literature about item design recommends certain practices for writing multiple-choice questions in order to increase item discrimination indices and increase the reliability of scores. Therefore, adherence to these principles while writing items is crucial for producing reliable scores. Implications of failure to adhere to these principles can unintentionally increase the difficulty of an item, decrease the discriminatory property of an item, and decrease the reliability of scores (Haladyna & Downing, 1989b). Because of these implications, studies that analyzed the results of voting on the clicker questions have likely analyzed unreliable scores, rendering their conclusions invalid. Studies that assessed students' affective beliefs about using clickers may have affected students' personal beliefs due to items being too difficult, confusing, or unfair.

Studies that assessed cognitive outcomes may have hindered students' learning from the clicker questions if the design was flawed. In addition, instructors who relied on the information they gathered from the voting results on flawed questions may have seen an inaccurate portrayal of students' understanding, possibly resulting in hindrance of student learning.

Educational Psychology Principles

Principles of educational psychology addressed in this review have been empirically researched and validated by educational psychologists for many decades; however, their application to use with clickers was a rarity in the literature. The following section outlines possibilities for future research in order to assess the implications of using principles, such as incorporating MERs, with clicker questions.

Since clickers are used in learning environments, it was expected that studies on the effectiveness of clickers would draw on research from educational psychology and related fields in order to support the theory, administration, item writing, or some other facet of the study. Only 41% of the articles analyzed in this review used references from highly cited journals that reference educational psychology principles. This is somewhat concerning as it indicates that educators using clickers in their classrooms are likely not employing empirically validated instructional methods. It is important for instructors, at all levels of learning, to employ sound methods proven to increase learning outcomes for their students.

Future Recommendations

Review of administrative procedures of studies on the cognitive and non-cognitive effects of clickers has shown incompatibilities among studies. These inconsistencies make any comparisons across studies difficult or invalid, especially if the studies possess extreme differences in administrative procedures. It is recommended that instructors considering

implementing clickers in their classrooms determine their primary desired outcome (cognitive or non-cognitive), administrative procedures they will use (e.g. re-voting, peer discussion), and consult studies that assessed the same outcomes using the same procedures.

The next steps from this review are to determine which administrative practices most consistently produced positive cognitive outcomes and to create a set of guidelines for clicker implementation based on those results. In addition, further research on the application of educational psychology principles to clicker implementation discussed in this paper is needed in order to determine the best practices for increasing learning using clickers in college classrooms. Specifically, effects on cognitive outcomes of using MERs in clicker questions as well as activating prior knowledge using clickers versus other methods should be explored. Based on current literature on learning gains facilitated by MERs and clickers, it is hypothesized that future research will show greater learning gains when used in conjunction. Based on research concerning non-cognitive outcomes of using clickers (e.g., increased student engagement, participation), it is also hypothesized that using clickers to activate prior knowledge will be more effective than other methods.

References

Addison, S., Wright, A., & Milner, R. (2009). Using clickers to improve student engagement and performance in an introductory biochemistry class. *Biochemistry and Molecular Biology Education, 37*(2), 84-91. doi:10.1002/bmb.20264

Adkins, D. C. (1960). *Test construction: Development and interpretation of achievement tests*. Columbus, Ohio: C. E. Merrill Books.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*(3), 183-198. doi:10.1016/j.learninstruc.2006.03.001

Airasian, P. W., & Miranda, H. (2002). The role of assessment in the revised taxonomy. *Theory into Practice, 41*(4), 249-254. doi:10.1207/s15430421tip4104_8

Alexander, P. A., Murphy, P. K., & Kulikowich, J. M. (1998). What responses to domain-specific analogy problems reveal about emerging competence: A new perspective on an old acquaintance. *Journal of Educational Psychology, 90*(3), 397-406. doi:10.1037/0022-0663.90.3.397

*Anderson, L. S., Healy, A. F., Kole, J. A., & Bourne, L. E. (2011). Conserving time in the classroom: The clicker technique. *The Quarterly Journal of Experimental Psychology*, *64*(8), 1457–1462.

*Anderson, L. S., Healy, A. F., Kole, J. A., & Bourne, L. E. (2013). The clicker technique: Cultivating efficient teaching and successful learning. *Applied Cognitive Psychology*, *27*(2), 222–234.

*Anthis, K. (2011). Is it the clicker, or is it the question? Untangling the effects of student response system use. *Teaching of Psychology*, *38*(3), 189–193.

*Auras, R., & Bix, L. (2007). WAKE UP! The effectiveness of a student response system in large packaging classes. *Packaging Technology and Science*, *20*(3), 183–195.

*Bandyopadhyay, A. (2013). Measuring the disparities between biology undergraduates' perceptions and their actual knowledge of scientific literature with clickers. *The Journal of Academic Librarianship*, *39*(2), 194–201.

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213-238. doi:10.2307/1170535

Beatty, I. D., Gerace, W. J., Leonard, W. J., & Dufresne, R. J. (2006). Designing effective questions for classroom response system teaching. *American Journal of Physics, 74*(1), 31-39. doi:10.1119/1.2121753

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74.

*Blasco-Arcas, L., Buil, I., Hernández-Ortega, B., & Sese, F. J. (2013). Using clickers in class. The role of interactivity, active collaborative learning and engagement in learning performance. *Computers & Education*, *62*, 102–110.

*Brady, M., Seli, H., & Rosenthal, J. (2013a). ‘‘Clickers’’ and metacognition: A quasi-experimental comparative study about metacognitive self-regulation and use of electronic feedback devices. *Computers & Education*, *65*, 56–63.

*Brady, M., Seli, H., & Rosenthal, J. (2013b). Metacognition and the influence of polling systems: How do clickers compare with low technology systems. *Educational Technology Research and Development*, *61*(6), 885–902.

*Bruck, A. D., & Towns, M. H. (2009). Analysis of classroom response system questions via four lenses in a general chemistry course. *Chemistry Education Research and Practice*, *10*(4), 291–295.

Butler, A., Godbole, N., & Marsh, E. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *Journal of Educational Psychology, 105*(2), 290-298. doi:10.1037/a0031026

*Campbell, C., & Monk, S. (2015). Introducing a learner response system to pre-service education students: Increasing student engagement. *Active Learning in Higher Education*, *16*(1), 25–36. doi:10.1177/1469787414558981

*Cardoso, W. (2011). Learning a foreign language with a learner response system: The students' perspective. *Computer Assisted Language Learning*, *24*(5), 393–417.

*Chen, J. C., Whittinghill, D. C., & Kadlowec, J. A. (2010). Classes that click: Fast, rich feedback to enhance student learning and satisfaction. *Journal of Engineering Education*, *99*(2), 159–168.

*Chen, T. L., & Lan, Y. L. (2013). Using a personal response system as an in-class assessment tool in the teaching of basic college chemistry. *Australasian Journal of Educational Technology*, *29*(1), 32–40.

*Cheung, G., Chan, K., Wan, K., & Ng, O. (2015). Understanding the adoption of a student response system from an integrated approach. In C. Watson (Ed.), *Proceedings of the 10th international conference on e-learning* (pp. 74–82). Bahamas, Nassau: Academic Conferences and Publishing International Limited.

*Chien, Y. T., Lee, Y. H., Li, T. Y., & Chang, C. Y. (2015). Examining the effects of displaying clicker voting results on high school students' voting behaviors, discussion processes, and learning outcomes. *Eurasia Journal of Mathematics Science and Technology Education*, *11*(5), 1089–1104.

*Clauson, K. A., Alkhateeb, F. M., & Singh-Franco, D. (2012). Concurrent use of an audience response system at a multi-campus college of pharmacy. *American Journal of Pharmaceutical Education, 76*(1), 6.

*Cotner, S. H., Fall, B. A., Wick, S. M., Walker, J. D., & Baepler, P. M. (2008). Rapid feedback assessment methods: Can we improve engagement and preparation for exams in large-enrollment courses? *Journal of Science Education and Technology*, *17*(5), 437–443.

*Crossgrove, K., & Curran, K. L. (2008). Using clickers in nonmajors- and majors-level biology courses: Student opinion, learning, and long-term retention of course material. *CBE-Life Sciences Education*, *7*(1), 146–154.

*Cruz, É., Dias, H., & Kortemeyer, G. (2011). The effect of formative assessment in Brazilian university physics courses. *Revista Brasileira de Ensino de Física*, *33*(4), 4315.

*Cubric, M., & Jefferies, A. (2015). The benefits and challenges of large-scale deployment of electronic voting systems: University student views from across different subject groups. *Computers & Education*, *87*, 98–111. doi:10.1016/j.compedu.2015.04.004

*Dallaire, D. H. (2011). Effective use of personal response ‘‘clicker’’ systems in psychology courses. *Teaching of Psychology*, *38*(3), 199–204.

*Deleo, P. A., Eichenholtz, S., & Sosin, A. A. (2009). Bridging the information literacy gap with clickers. *The Journal of Academic Librarianship*, *35*(5), 438–444.

*Desrochers, M. N., & Shelnutt, J. M. (2012). Effect of answer format and review method on college students' learning. *Computers & Education*, *59*(3), 946–951.

*Dill, E. (2008). Do clickers improve library instruction? Lock in your answers now. *The Journal of Academic Librarianship*, *34*(6), 527–529.

*Duret, D., & Senior, A. (2015). Comparative study of three different personal response systems with fourth-year undergraduate veterinary students. *Journal of Veterinary Medical Education*, *42*(2), 120–126. doi:10.3138/jvme.0814–079R2

*Elicker, J. D., & McConnell, N. L. (2011). Interactive learning in the classroom: Is student response method related to performance? *Teaching of Psychology*, *38*(3), 147–150.

*Fallon, M., & Forrest, S. L. (2011). High-tech versus low-tech instructional strategies: A comparison of clickers and handheld response cards. *Teaching of Psychology*, 38(3), 194–198.

*FitzPatrick, K. A., Finn, K. E., & Campisi, J. (2011). Effect of personal response systems on student perception and academic performance in courses in a health sciences curriculum. *Advances in Physiology Education*, *35*(3), 280–289.

*Flynn, A. B. (2011). Developing problem-solving skills through retrosynthetic analysis and clickers in organic chemistry. *Journal of Chemical Education*, *88*(11), 1496–1500.

*Flynn, A. B. (2012). Development of an online, post class question method and its integration with teaching strategies. *Journal of Chemical Education*, *89*(4), 456–464.

*Fortner-Wood, C., Armistead, L., Marchand, A., & Morris, F. B. (2013). The effects of student response systems on student learning and attitudes in undergraduate psychology courses. *Teaching of Psychology*, *40*(1), 26–30.

*Freeman, S., Haak, D., & Wenderoth, M. P. (2011). Increased course structure improves performance in introductory biology. *CBE-Life Sciences Education*, *10*(2), 175–186.

*Friedline, T., Mann, A. R., & Lieberman, A. (2013). Teaching note—Ask the audience: Using student response systems in social work education. *Journal of Social Work Education*, *49*(4), 782–792.

*Gauci, S. A., Dantas, A. M., Williams, D. A., & Kemm, R. E. (2009). Promoting student-centered active learning in lectures with a personal response system. *Advances in Physiology Education*, *33*(1), 60–71.

*Gebru, M. T., Phelps, A. J., & Wulfsberg, G. (2012). Effect of clickers versus online homework on students' long-term retention of general chemistry course material. *Chemistry Education Research and Practice*, *13*(3), 325–329.

Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco: W. H. Freeman.

Gierl, M. J., Bulut, O., Guo, Q., & Zhang, X. (2017). Developing, analyzing, and using distractors for multiple-choice tests in education: A comprehensive review. *Review of Educational Research, 87*(6), 1082-1116. doi:10.3102/0034654317726529

*Gök, T. (2011a). An evaluation of student response systems from the viewpoint of instructors and students. *The Turkish Online Journal of Educational Technology*, *10*(4), 67–83.

*Gök, T. (2011b). Using the classroom response system to enhance students' learning and classroom interactivity. *Eurasian Journal of Educational Research*, *45*, 49–68.

*Goldberg, F., Price, E., Robinson, S., Boyd-Harlow, D., & McKean, M. (2012). Developing the learning physical science curriculum: Adapting a small enrollment, laboratory and discussion based physical science course for large enrollments. *Physical Review Special Topics-Physics Education Research*, *8*(1), 010121.

*Gray, K., Owens, K., Liang, X., & Steer, D. (2012). Assessing multimedia influences on student responses using a personal response system. *Journal of Science Education and Technology*, *21*(3), 392–402.

*Grzeskowiak, L. E., Thomas, A. E., To, J., Reeve, E., & Phillips, A. J. (2015). Enhancing continuing education activities using audience response systems: A single-blind controlled trial. *Journal of Continuing Education in the Health Professions*, *35*(1), 38–45. doi:10.1002/chp.21269

Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Boston: Allyn and Bacon.

Haladyna, T. M., & Downing, S. M. (1989a). A taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education, 2*(1), 37-50. doi:10.1207/s15324818ame0201_3

Haladyna, T. M., & Downing, S. M. (1989b). Validity of a taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education, 2*(1), 51-78. doi:10.1207/s15324818ame0201_4

*Han, J. H. (2014). Unpacking and repacking the factors affecting students' perceptions of the use of classroom communication systems (CCS) technology. *Computers & Education*, *79*, 159–176.

*Han, J. H., & Finkelstein, A. (2013). Understanding the effects of professors' pedagogical development with clicker assessment and feedback technologies and the impact on students' engagement and learning in higher education. *Computers & Education*, *65*, 64–76.

*Hoekstra, A. (2015). Because you don't realize how many people have different experiences than you: Effects of clicker use for class discussions in sociology. *Teaching Sociology*, *43*(1), 53–60. doi:10.1177/0092055x14556054

Hunsu, N., Adesope, O., & Bayly, D. (2016). A meta-analysis of the effects of audience response systems (clicker-based technologies) on cognition and affect. *Computers & Education, 94*, 102-119. doi:10.1016/j.compedu.2015.11.013

Kay, R. H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education, 53*(3), 819-827. doi:10.1016/j.compedu.2009.05.001

*Kim, Y., Jeong, S., Ji, Y., Lee, S., Kwon, K. H., & Jeon, J. W. (2015). Smartphone response system using twitter to enable effective interaction and improve engagement in large classrooms. *IEEE Transactions on Education*, *58*(2), 98–103. doi:10.1109/te.2014.2329651

*Knight, J. K., Wise, S. B., Rentsch, J., & Furtak, E. M. (2015). Cues matter: Learning assistants influence introductory biology student interactions during clicker-question discussions. *CBE-Life Sciences Education*, *14*(4). doi:10.1187/cbe.15–04-0093

Kostons, D., & van der Werf, M. (2015). The effects of activating prior topic and metacognitive knowledge on text comprehension scores. *British Journal of Educational Psychology, 85*(3), 264-275. doi:10.1111/bjep.12069

Krathwohl, D. R. (2002). A revision of bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212-218. doi:10.1207/s15430421tip4104_2

*Kvadsheim, R., Haugerud, H., Hammer, H. L., Bratterud, A., & Habib, L. (2015). Does clicker use improve exam scores? A controlled randomized experiment in a bachelor-level course in software engineering. *International Journal of Engineering Education*, *31*(2), 505–520.

*Landrum, R. E. (2013). The ubiquitous clicker SoTL applications for scientist-educators. *Teaching of Psychology*, *40*(2), 98–103.

*Lee, H., & Feldman, A. (2015). Photographs and classroom response systems in middle school astronomy classes. *Journal of Science Education and Technology*, *24*(4), 496–508. doi:10.1007/s10956–014-9539-z

*Lee, H., & Schneider, S. E. (2015). Using astronomical photographs to investigate misconceptions about galaxies and spectra: Question development for clicker use. *Physical Review Special Topics-Physics Education Research*, *11*(2). doi:10.1103/PhysRevSTPER.11.020101

*Lee, U. J., Sbeglia, G. C., Ha, M., Finch, S. J., & Nehm, R. H. (2015). Clicker score trajectories and concept inventory scores as predictors for early warning systems for large stem classes. *Journal of Science Education and Technology*, *24*(6), 848–860. doi:10.1007/s10956–015-9568–2

*Levesque, A. A. (2011). Using clickers to facilitate development of problem-solving skills. *CBE-Life Sciences Education*, *10*(4), 406–417.

*Lin, Y. C., Liu, T. C., & Chu, C. C. (2011). Implementing clickers to assist learning in science lectures: The clicker-assisted conceptual change model. *Australasian Journal of Educational Technology*, *27*(6), 979–996.

Liu, C., Chen, S., Chi, C., Chien, K., Liu, Y., & Chou, T. (2017). The effects of clickers with different teaching strategies. *Journal of Educational Computing Research, 55*(5), 603-628.

*Loukomies, A., Juuti, K., & Lavonen, J. (2015). Investigating situational interest in primary science lessons. *International Journal of Science Education*, *37*(18), 3015–3037. doi:10.1080/09500693.2015.1119909

*Ludvigsen, K., Krumsvik, R., & Furnes, B. (2015). Creating formative feedback spaces in large lectures. *Computers & Education*, *88*, 48–63. doi:10.1016/j.compedu.2015.04.002

*Lundeberg, M. A., Kang, H., Wolter, B., delMas, R., Armstrong, N., Borsari, B., . . . Herreid, C. F. (2011). Context matters: Increasing understanding with interactive clicker case studies. *Educational Technology Research and Development*, *59*(5), 645–671.

*MacArthur, J. R., & Jones, L. (2013). Self-assembled student interactions in undergraduate general chemistry clicker classrooms. *Journal of Chemical Education*, *90*(12), 1586–1589.

*MacGeorge, E. L., Homan, S. R., Dunning, J. B. Jr., Elmore, D., Bodie, G. D., Evans, E., . . . Geddes, B. (2008). Student evaluation of audience response technology in large lecture classes. *Educational Technology Research and Development*, *56*, 125–145.

Macmillan Learning (2017). iClicker. Retrieved from https://www.iclicker.com

*Marlow, D. W. (2010). Engaging syntax: Using a personal response system to encourage grammatical thought. *American Speech*, *85*(2), 225–237.

Mayer, R. E. (1979). Twenty years of research on advance organizers: Assimilation theory is still the best predictor of results. *Instructional Science, 8*(2), 133-167. doi:10.1007/BF00117008

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, *32*(1), 1-19.

*Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., . . . Zhang, H. J. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, *34*(1), 51–57.

*McDonough, K., & Foote, J. A. (2015). The impact of individual and shared clicker use on students' collaborative learning. *Computers & Education*, *86*, 236–249. doi:10.1016/j.compedu.2015.08.009

*McLaughlin, J. E., Gharkholonarehe, N., Khanova, J., Deyo, Z. M., & Rodgers, J. E. (2015). The impact of blended learning on student performance in a cardiovascular pharmacotherapy course. *American Journal of Pharmaceutical Education*, *79*(2), 24.

*Medina, M. S., Medina, P. J., Wanzer, D. S., Wilson, J. E., Er, N., & Britton, M. L. (2008). Use of an audience response system (ARS) in a dual-campus classroom environment. *American Journal of Pharmaceutical Education*, *72*(2), 38.

*Miller, K., Schell, J., Ho, A., Lukoff, B., & Mazur, E. (2015). Response switching and self-efficacy in peer instruction classrooms. *Physical Review Special Topics—Physics Education Research*, *11*(1). doi:10.1103/PhysRevSTPER.11.010104

*Mollborn, S., & Hoekstra, A. (2010). ‘‘A meeting of minds’’ using clickers for critical thinking and discussion in large sociology classes. *Teaching Sociology*, *38*(1), 18–27.

*Morice, J., Michinov, N., Delaval, M., Sideridou, A., & Ferrieres, V. (2015). Comparing the effectiveness of peer instruction to individual learning during a chromatography course. *Journal of Computer Assisted Learning*, *31*(6), 722–733. doi:10.1111/jcal.12116

*Morling, B., McAuliffe, M., Cohen, L., & DiLorenzo, T. M. (2008). Efficacy of personal response systems (‘‘clickers’’) in large, introductory psychology classes. *Teaching of Psychology*, *35*(1), 45–50.

*Moss, K., & Crowley, M. (2011). Effective learning in science: The use of personal response systems with a wide range of audiences. *Computers & Education*, *56*(1), 36–43.

*Murphy, K. (2012). Using a personal response system to map cognitive efficiency and gain insight into a proposed learning progression in preparatory chemistry. *Journal of Chemical Education*, *89*(10), 1229–1235.

*Muthyala, R. S., & Wei, W. (2013). Does space matter? Impact of classroom space on student learning in an organic-first curriculum. *Journal of Chemical Education*, *90*(1), 45–50.

*Pagano, R., & Paucar-Caceres, A. (2013). Using systems thinking to evaluate formative feedback in UK higher education: The case of classroom response technology. *Innovations in Education and Teaching International*, *50*(1), 94–103.

*Patterson, B., Kilpatrick, J., & Woebkenberg, E. (2010). Evidence for teaching practice: The impact of clickers in a large classroom environment. *Nurse Education Today*, *30*(7), 603–607.

*Perez, K. E., Strauss, E. A., Downey, N., Galbraith, A., Jeanne, R., & Cooper, S. (2010). Does displaying the class results affect student discussion during peer instruction? *CBE-Life Sciences Education*, *9*(2), 133–140.

*Pettit, R. K., McCoy, L., Kinney, M., & Schwartz, F. N. (2015). Student perceptions of gamified audience response system interactions in large group lectures and via lecture capture technology. *BMC Medical Education*, *15*(1), 92. doi:10.1186/s12909–015-0373–7

Popham, W. J. (2011). *Classroom assessment: What teachers need to know*. Boston, MA: Allyn & Bacon.

*Prunuske, A. J., Batzli, J., Howell, E., & Miller, S. (2012). Using online lectures to make time for active learning. *Genetics*, *192*(1), 67–72.

*Ross, A., & Furno, C. (2011). Active learning in the library instruction environment: An exploratory study. *Portal: Libraries and the Academy*, *11*(4), 953–970.

*Russell, J. S., McWilliams, M., Chasen, L., & Farley, J. (2011). Using clickers for clinical reasoning and problem solving. *Nurse Educator*, *36*(1), 13–15.

*Ryan, B. J. (2013). Line up, line up: Using technology to align and enhance peer learning and assessment in a student centred foundation organic chemistry module. *Chemistry Education Research and Practice*, *14*(3), 229–238.

*Satheesh, K. M., Saylor-Boles, C. D., Rapley, J. W., Liu, Y., & Gadbury-Amyot, C. C. (2013). Student evaluation of clickers in a combined dental and dental hygiene periodontology course. *Journal of Dental Education*, *77*(10), 1321–1329.

*Schuller, M. C., DaRosa, D. A., & Crandall, M. L. (2015). Using just-in-time teaching and peer instruction in a residency program's core curriculum: Enhancing satisfaction, engagement, and retention. *Academic Medicine*, *90*(3), 384–391. doi:10.1097/acm.00 00000000000578

*Schullo-Feulner, A., Kolar, C., & Janke, K. K. (2015). A five-year evaluation of examination structure in a cardiovascular pharmacotherapy course. *American Journal of Pharmaceutical Education*, *79*(7), 98.

*Shaffer, D. M., & Collura, M. J. (2009). Evaluating the effectiveness of a personal response system in the classroom. *Teaching of Psychology*, *36*(4), 273–277.

*Shapiro, A. M., & Gordon, L. T. (2012). A controlled study of clicker-assisted memory enhancement in college classrooms. *Applied Cognitive Psychology*, *26*(4), 635–643.

*Shieh, R. S., & Chang, W. (2013). Implementing the interactive response system in a high school physics context: Intervention and reflections. *Australasian Journal of Educational Technology*, *29*(5), 748–761.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189. doi:10.3102/0034654307313795

*Smith, D. A., & Rosenkoetter, M. M. (2009). Effectiveness, challenges, and perceptions of classroom participation systems. *Nurse Educator*, *34*(4), 156–161.

*Smith, M. K., Annis, S. L., Kaplan, J. J., & Drummond, F. (2012). Using peer discussion facilitated by clicker questions in an informal education setting: Enhancing farmer learning of science. *PloS One*, *7*(10), e47564.

*Smith, M. K., Trujillo, C., & Su, T. T. (2011). The benefits of using clickers in small-enrollment seminar-style biology courses. *CBE-Life Sciences Education*, *10*(1), 14–17.

*Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science*, *323*(5910), 122–124.

*Smith, M. K., Wood, W. B., Krauter, K., & Knight, J. K. (2011). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE-Life Sciences Education*, *10*(1), 55–63.

*Stehling, V., Schuster, K., Richert, A., & Isenhardt, I. (2015). Please vote now! Evaluation of audience response systems - First results from a flipped classroom setting. In C. Watson (Ed.), *Proceedings of the 10th international conference on e-learning* (pp. 287–295). Bahamas, Nassau: Academic Conferences and Publishing International Limited.

Stiggins, R. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *The Phi Delta Kappan, 87*(4), 324-328. doi:10.1177/003172170508700414

*Stowell, J. R. (2015). Use of clickers vs. mobile devices for classroom polling. *Computers & Education*, *82*, 329–334. doi:10.1016/j.compedu.2014.12.008

*Stowell, J. R., & Nelson, J. M. (2007). Benefits of electronic audience response systems on student participation, learning, and emotion. *Teaching of Psychology*, *34*(4), 253–258.

*Stowell, J. R., Oldham, T., & Bennett, D. (2010). Using student response systems (''clickers'') to combat conformity and shyness. *Teaching of Psychology*, *37*(2), 135–140.

*Such, J. M., Criado, N., & Garcia-Fornes, A. (2015). An active learning technique enhanced with electronic polls. *International Journal of Engineering Education*, *31*(4), 1048–1057.

*Sutherlin, A. L., Sutherlin, G. R., & Akpanudo, U. M. (2013). The effect of clickers in university science courses. *Journal of Science Education and Technology*, *22*(5), 651–666.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251-296.

Turning Technologies, LLC (2017). TurningPoint. Retrieved from https://www.turningtechnologies.com/turningpoint

*Vana, K. D., Silva, G. E., Muzyka, D., & Hirani, L. M. (2011). Effectiveness of an audience response system in teaching pharmacology to baccalaureate nursing students. *Computers, Informatics, Nursing*, *29*, 105–113.

*Velasco, M., & Cavdar, G. (2013). Teaching large classes with clickers: Results from a teaching experiment in comparative politics. *Political Science & Politics*, *46*, 823–829.

*Vital, F. (2011). Creating a positive learning environment with the use of clickers in a high school chemistry classroom. *Journal of Chemical Education*, *89*(4), 470–473.

*Voelkel, S., & Bennett, D. (2014). New uses for a familiar technology: Introducing mobile phone polling in large classes. *Innovations in Education and Teaching International*, *51*(1), 46–58.

*Waltz, E. C., Maniccia, D. M., Bryde, R. L., Murphy, K., Harris, B. R., & Waldenmaier, M. N. (2010). Training the public health workforce from Albany to Zambia: Technology lessons learned along the way. *Public Health Reports*, *125*, 61–69.

*Wang, A. I. (2015). The wear out effect of a game-based student response system. *Computers & Education*, *82*, 217–227. doi:10.1016/j.compedu.2014.11.004

*Warnich, P., & Gordon, C. (2015). The integration of cell phone technology and poll everywhere as teaching and learning tools into the school history classroom. *Yesterday and Today*, *13*, 40–66.

*Willoughby, S. D., & Gustafson, E. (2009). Technology talks: Clickers and grading incentive in the large lecture hall. *American Journal of Physics*, *77*(2), 180–183.

Wittrock, M. C. (1990). Generative processes of comprehension. *Educational Psychologist*, *24*, 354–376.

*Yu, Z., Chen, W., Kong, Y., Sun, X. L., & Zheng, J. (2014). The impact of clickers instruction on cognitive loads and listening and speaking skills in college English class. *PloS ONE*, *9*(9), e106626.

*Zingaro, D. (2015). Examining interest and grades in computer science 1: A study of pedagogy and achievement goals. *ACM Transactions on Computing Education*, *15*(3). doi:10.1145/2802752

Appendix: Table of Journals Referencing Educational Psychology Principles

Category	Journal Name	SJR	Rank #
Educational Psychology	*Educational Psychologist*	4.760	1
	Cognitive Psychology	3.380	4
	Learning and Instruction	2.803	9
	Cognition	2.778	11
	Journal of Educational Psychology	2.467	16
	Journal of the Learning Sciences	2.080	21
	Contemporary Educational Psychology	1.898	30
	Educational Psychology Review	1.645	38
	Cognition and Instruction	1.470	42
	Instructional Science	1.323	55
Assessment	*Journal of Educational Measurement*	2.752	12
	Psychometric Bulletin and Review	2.044	24
	Applied Psychological Measurement	1.420	28*
	Educational and Psychological Measurement	1.445	45

Note. Rank # reflects the rank of the journal in the category of Educational Psychology unless otherwise indicated.

**Applied Psychological Measurement* was ranked number 28 in the category of Miscellaneous Psychology