University of Nevada, Reno

Reducing Racing During Online Instruction with Contingent Postfeedback Delays

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in
Psychology

by

Erick Mark Dubuque

May 2012

ABSTRACT

Online instruction has become more prevalent in higher education over the last decade. Research investigating the effectiveness of this mode of instruction indicates that it is as good as, or better than, instruction in traditional classroom settings. In addition to the various benefits of online learning, there are also unique challenges that accompany this form of instruction. One of these challenges is racing during self-paced lessons. When racing occurs students are not attending to the instructional material and are instead moving through a lesson in an apparent attempt to escape the instructional context as quickly as possible. The purpose of our study was to assess the impact of contingent postfeedback delays on racing during an online lesson. Results from three separate experiments indicate that postfeedback delays can be an effective way to manage racing during a self-paced online lesson for some students.

DESCRIPTORS: racing, online learning, postfeedback delays

ACKNOWLEDGEMENTS

A number of people deserve recognition for supporting me during my academic career and helping me to come this far. I would like to say thank you to my parents for always encouraging me to pursue my goals and for teaching me that nothing is beyond my reach. I would also like to recognize my wife, Molly. Her love, support, and contributions to all of my ideas have allowed me to become the person I am today. I owe a special thank you to Roger Tudor for introducing me to the science of behavior and for helping me see to it that I ended up in the graduate program of my choosing. The faculty, my colleagues and friends in the University of Nevada, Reno Behavior Analysis Program deserve my appreciation as well for always challenging me to do better. This is especially true of Ramona Houmanfar who guided and supported me during the early years of my graduate training and to my lab mates who never hesitated to provide their valuable feedback during my presentations. No one has been a bigger influence on my career as a behavior scientist than Linda Hayes. She has been a mentor, close friend, and an inspiration. I look up to her more than she will probably ever know. Finally, I would like to say thank you to my son Jack. Being a father to my son is the greatest role I have ever had. A role that makes me strive harder to be better in everything I do.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Online learning has become a pervasive mode of instruction within higher education over the last decade (Allen & Seaman, 2011). This change in instructional format represents a significant shift from traditional face-to-face interactions resulting in learning opportunities that cannot be replicated in the classroom (Fish & Wickersham, 2009; Johnson & Rubin, 2011). One of the primary benefits of online instruction over classroom instruction is the ability to construct lessons that are student-paced as opposed to instructor-paced (Johnson & Dickinson, in press). In other words, instruction taking place in an online context can be easily tailored so students can move through lesson content at their own pace. In contrast, instruction taking place in a classroom setting is almost always instructor-paced. This means that lesson content is presented to students at the same rate regardless of their proficiency with the material being reviewed. In an online context, individualizing instruction so students have the ability to control how quickly they move through lesson content can save time for fluent performers and result in better learning outcomes for slower performers who need more time with instructional material. Although it is generally considered beneficial to allow students to move through instructional content at their own pace this arrangement does not always result in optimal performance. When a learner moves quickly and inaccurately through a self-paced lesson it is known as "racing" (Crosbie & Kelly, 1993, 1994; Johnston & Dickinson, in press; Kelly & Crosbie, 1997; Munson & Crosbie, 1998). Racing is presumed to occur so learners can contact the conditioned reinforcer of being finished with instructional content more quickly. The purpose of this study is to address this contingency directly by assessing the impact of contingent postfeedback delays on racing during an online lesson.

The Growth of Online Education

Over the last decade online instruction has become ubiquitous in higher education. In 2010, enrollment in online courses represented 30% of total enrollment based on survey results collected from over

2,500 colleges and universities (Allen & Seaman, 2011). In the same survey, over 65% of Chief Academic Officers reported that online education was critical to the long-term strategy of their institutions. In 2002, around 1.6 million higher education students had taken at least 1 online course; in 2010, this number almost quadrupled to over 6 million (Allen & Seaman, 2011).

There are numerous reasons that account for the quick rise of online education and not all of them are purely academic. Many students consider online courses more convenient, flexible, and accommodating to busy schedules. Offering educational opportunities using an online format allows universities and colleges to reach a wider market. Additionally, online education is considered more environmentally friendly than traditional modes of instruction, as hosting materials on a student's computer reduces the need for paper and fuel to drive to physical classrooms.

Online Education is a Viable Method of Instruction

Regardless of the non-academic benefits, given the widespread demand and adoption of online education over the last decade it is clearly important to establish how this form of instruction impacts learning. Fortunately, the research on computer-based instruction (CBI), and more specifically online instruction, find that these methods are viable, and in many cases better alternatives, than face-to-face instruction. Johnson & Rubin (2011) conducted an extensive review of the interactive (i.e. requiring student responding) CBI literature and found that an overwhelming 95.2% of the time, interactive CBI used alone or as a supplement was as good or better than other instructional alternatives. In addition, results from a meta-analysis of online education studies conducted by the U.S. Department of Education found that students enrolled in an online or hybrid course (online & face-to-face) performed better than their peers receiving solely face-to-face instruction (Means, Toyama, Murphy, Bakia & Jones, 2009).

With such wide variation in how content is delivered online it is difficult to parse out the reasons why this form of instruction can lead to better learning gains. Means et al. (2009) claim that online learning isn't necessarily a superior medium over face-to-face instruction. Instead they surmise that advantages documented in online instruction may be attributed to other factors, including increased exposure to learning materials. There is some evidence to support the claim that additional practice leads to better learning outcomes in CBI (Schnackenberg & Sullivan, 2000; Shute, Gawlick & Gluck, 1998). For example, Green, Eppler, Ironsmith & Wuensch (2007) found that using branching scenarios which exposed learners to additional instructional materials to remediate errors in a CBI lesson resulted in better learning gains than using a linear instructional sequence. Additionally, several studies have shown that simply inserting a 5-10 second postfeedback delay between questions in an CBI lesson results in better performance (Crosbie & Kelly, 1994; Johnson & Dickinson, in press; Kelly & Crosbie, 1997). Finally, using CBI to enforce a mastery criterion before allowing the learner to move on has been shown to be more effective than linear instruction where fewer learning trials are normally required (Montazemi & Wang, 1995; Moreno & Valdez, 2005).

Even if additional contact with instructional content is the primary means by which better learning gains occur, CBI offers opportunities for programming instruction that are simply not feasible using face-to-face instruction. In this respect, CBI, and more specifically online instruction, have nothing to gain by trying to mimic traditional classroom instruction (Fish & Wickersham, 2009; Johnson & Rubin, 2011). Instead educators working in these settings are better off utilizing instructional strategies that take specific advantage of the format. For example, one of the biggest advantages CBI has over face-to-face instruction is the ability to tailor instruction to the individual leaner. Research shows that CBI which capitalizes on this advantage produces better performance. For instance, in addition to the strategies mentioned previously, CBI lessons which incorporate contingent (Morrison, Ross, Gopalakrishnan & Casey, 1995), expert (Gibbons, Robertson & Thompson, 2001) and/or elaborate (Chase & Houmanfar, 2009) feedback based on individual learner responses are known to produce better outcomes.

Another advantage of CBI over traditional face-to-face instruction is the ability to require a response from every learner when a question is asked. Studies show that requiring learners to make overt responses during CBI leads to better performance (Bodemer et al., 2004; Miller & Malott, 1997; Tudor, 1995). This is especially true when the required responses involve composing an answer (e.g., fill-in-the-blank question) as opposed to selecting an answer (e.g., multiple-choice questions) from a set of choices (Clariana, 2003, 2004; Katayama, Shambaugh & Doctor, 2005). Finally, CBI lessons which allow for multiple pathways to a correct answer also appear to promote better performance (Yadrick, Regian, Connolly-Gomez & Roberston-Schule, 1997).

New Mode of Instruction, New Challenges

While there are obvious advantages of CBI over face-to-face instruction there are also some very clear challenges, especially to online instruction, that need to be addressed. For example, despite evidence to the contrary, most Chief Academic Officers and students in institutions of higher learning still do not accept the value and legitimacy of online education and perceive this form of instruction as leading to inferior learning outcomes (Allen & Seaman, 2010; Young & Norgard, 2006). A contributing factor to this perception is likely to be the large variation in the quality of online instruction currently being employed (Hewett & Powers, 2007). This variation can probably be attributed to a lack in the technical expertise that is required of an instructor to develop online instructional content which leads to significant learning gains. Unfortunately, based on survey results from Allen (2010), about one-fifth of higher education institutions offer no training to their instructors on teaching courses online. Further, of those institutions that do offer such training, 59% rely on informal mentoring. Finally, developing quality content designed to implement all of the elements shown to lead to better learning outcomes is both extremely time and resource intensive on the front-end, especially when compared to the lecture-based format used in traditional classrooms. Chapman (2010) surveyed 249 companies representing almost 4,000 learning development professionals. Respondents were asked how long it takes to develop 1-hour of instructional materials across various modalities. On average, respondents reported that it takes 43 hours for instructor-led training (i.e., face-to-face instruction); 79 hours for a basic e-learning (e.g., simple text graphics, some audio, test questions, etc.); 184 hours for intermediate e-learning (e.g., use of animations and multimedia, etc.); and 490 hours for advanced e-learning (simulations, custom interactions, games, etc.). In addition, these same respondents were asked how much it costs to develop 1-hour of instructional materials across the same modalities. On average, respondents reported it costs $5,934 for instructor-led training; $10,054 for basic e-learning, $18,583 for interactive e-learning, and $50,371 for advanced e-learning. Given the relative novelty of online education, combined with the time, cost, and expertise involved in creating quality online instructional materials, it is not surprising that so much variation can be found.

Racing

In addition, development time, costs, and training, online learning also presents instructors with other challenges not found in the traditional classroom. One of these challenges involves teaching students in a context where competing activities vie for their attention. When a student takes classes in a physical classroom they are limited from contacting outside activities that might otherwise compete for their attention. However, when students are completing lessons in an online format outside of the classroom they often have access to competing activities, whereby it is impossible for an instructor to determine whether they are fully engaged with the instructional tasks presented to them. Combining easy access to competing activities with a self-paced lesson sets the conditions for racing to occur. When racing occurs students are not attending to the instructional material and are instead moving through a lesson in an apparent attempt to escape the instructional context as quickly as possible. Assuming that racing is controlled by escape from aversive instructional content, it is logical to further presume that racing is more likely to occur when lessons take place in the presence of competing activities that may function as reinforcers for escape. In other words, students taking an online lesson are more likely to be in the presence of competing activities (e.g., email, Facebook®, television, etc.) that could function as reinforcers for escaping or rushing through aversive instructional content.

Thus far racing has been empirically studied and observed only in computer-based laboratory settings. Only two methods have been shown to be effective at reducing racing during self-paced instruction in a computer-based setting. The first method involves using monetary incentives and disincentives for correct and incorrect responding. Munson & Crosbie (1998) used this method and showed a 10% gain in performance after giving and taking away $0.05 following correct and incorrect answers in a computer-based lesson. However, Johnson & Dickinson (in press) used a similar experimental preparation and found monetary incentives and disincentives did not have a significant impact on performance. These mixed results suggest that further research may be needed to determine the effectiveness of monetary incentives and disincentives on racing during computer-based, self-paced instruction. Furthermore, there are limitations to using monetary incentives that prevent practical adoption. Even though Munson & Crosbie demonstrated that incentives as low as $0.05 can be effective, these costs can add up quickly, especially when multiplied across every lesson over an entire semester and across all students in a class. For example, assuming all answers are correct, a $0.05 incentive for 20 questions every week across a 15-week semester for 35 students equates to $525.00 for a single class in a semester.

Postfeedback Delays

Postfeedback delays are the second method that has been shown to be effective at reducing racing (Crosbie & Kelly, 1993, 1994; Kelly & Crosbie, 1997) during CBI. Postfeedback delays occur when an enforced delay follows performance feedback in a computer-based lesson. Researchers began studying the impact of intertrial interval delays on human performance in the mid-20th century (Brown, 1949; Bilodeau & Bilodeau, 1958). Early experiments in this area were concerned primarily with investigating the effect of delays between a subject's response and informative feedback (i.e. pre-feedback delays). Bilodeau &

Bilodeau (1958) were the first to specifically examine postfeedback delays. They reported that extraordinarily long postfeedback delays of 24 hours and 1 week had a deleterious effect on the performance of subjects attempting to learn a simple task with unspecified rules (i.e., pulling a yardstick 10 inches out of a sheath while blindfolded). Results from Denny, Marvel, Hall & Rokeach (1960) confirmed Bilodeau & Bilodeau's (1958) findings using significantly shorter delay periods (20 seconds) and further identified unspecified or unclear rules given to subjects (e.g., "Draw a line 60 glubs long.") as a crucial variable required to observe the negative impact that postfeedback delays could have on performance. Findings from later studies (Bourne & Bunderson, 1963; White & Schmidt, 1972) challenged the idea that postfeedback delays were harmful to performance by investigating their impact in the context of a concept identification task. Rather than finding a decrease in performance for subjects in the delay condition, these researchers discovered that performance increases were positively correlated with longer postfeedback delays. They also found that when complexity of the concept identification task was increased, postfeedback delays played an even greater role in improving performance.

It wasn't until 1994, that Crosbie and Kelly specifically examined the impact of postfeedback delays on racing during a computerized self-paced lesson. In their first experiment, Crosbie & Kelly (1994) exposed 4 subjects to an instructional sequence adapted from Holland & Skinner's (1961) programed text, *Analysis of Behavior*, into a computerized format. The researchers were interested in investigating the impact of postfeedback delays on performance across 3 conditions. These conditions included: (a) no delay between questions regardless of answer accuracy, (b) a 10-second delay after each question regardless of answer accuracy (i.e., non-contingent delay), and (c) a 10-second delay following questions answered incorrectly (i.e., response contingent delay). Additionally, subjects could view the instructional content during the pauses between trials in the non-contingent and contingent delay conditions. Results from this study showed no substantial difference in responding between the no delay and contingent delay conditions. However, the researchers observed that subjects performed better during the non-contingent delay condition when compared to the other two conditions. Based on these data, the authors dismissed the 10-second time delay as an effective punisher (i.e., time-out procedure) and instead surmised that the extra time spent looking at the content during the 10-second pause after each question may have resulted in better performance.

This was a reasonable assumption given that most of the instructional content used in their lesson was built off of earlier content presented to their subjects. To further determine whether better performance was due to an increase in contact with instructional material or whether it was due to preventing the subject from racing (i.e., pacing the subject's exposure to the material) the authors ran another experiment (Crosbie & Kelly, 1994, Experiment 2). Once again, they investigated how the performance of 4 subjects was impacted under 3 conditions but with a new modification. The conditions for the new experiment included: (a) no delay between questions regardless of answer accuracy, (b) a 10-second delay after each question regardless of answer accuracy (non-contingent delay) and (c) a 10-second delay after each question accompanied by a blank screen while the program was paused (non-contingent blank). Results from this second experiment provided evidence that the extra time subjects had to examine the instructional content (i.e., non-contingent delay) was likely responsible for the better performance over the other two conditions, no delay and non-contingent delay accompanied by a blank screen.

In 1997, Kelly & Crosbie followed-up their investigations in this area with another experiment designed to refine their previous experimental design. The authors modified the previous experiment (Crosbie & Kelly, 1994, Experiment 2) by shortening the experimental sessions, adding time-based measures (although the results of these measures were only vaguely mentioned), and added pre-, post-, and follow-up tests to assess the impact of their interventions on performance prior, directly after, and at a later date. Results from this study continued to support the finding that allowing an opportunity to review instructional content during a forced postfeedback delay results in better performance. A later study conducted by Johnson

& Dickinson (in press) confirmed this finding even with postfeedback delays as short as 5-seconds.

Although the studies on postfeedback delays provide some important contributions to the research investigating racing during CBI, there are some limitations that prevent these findings from generalizing to an online lesson. Most importantly, all of the research investigating postfeedback delays during CBI (Crosbie & Kelly, 1994; Johnson & Dickinson, in press; Kelly & Crosbie, 1997) has examined this intervention in conjunction with exposure to extra instructional time. This means that even though the authors of these studies describe their interventions as postfeedback delays, this label fails to adequately acknowledge the most important aspect of their intervention; namely the extra exposure to instructional content during the delay. Continuing to refer to this intervention simply as postfeedback delay when the only study investigating the impact of the delay without the extra instructional time (Crosbie & Kelly, 1994) was demonstrated to be ineffective is misleading. Further, even though the results from Crosbie & Kelly (1994) showed that 10-second postfeedback delays (without exposure to instructional content) had no significant impact on academic performance, other considerations suggest caution in generalizing this finding to self-paced performance in an online context.

First, the Crosbie & Kelly (1994) study and all other studies (Johnson & Dickinson, in press; Kelly & Crosbie, 1997) investigating postfeedback delays with extra instructional exposure were conducted in highly controlled laboratory settings far removed from most real-life competing activities (e.g., watching television, surfing the web, chatting with friends, eating food, etc.) that could have had an impact on performance. In other words, the subjects in these experiments probably had nothing better to do than read what was on the screen in front of them while waiting for the 5-10 second pause to end. It is not clear whether this would still be the case if the experiments had been conducted on a computer with internet access where subjects could have checked their email or looked up sports scores during the delay. Second, it is possible that running the study in a controlled laboratory setting may have set the conditions for subjects to perform better knowing that the experimenters were present in the next room. Third, when a lesson is taken online it is completed in a physical location of the student's choosing (e.g., bedroom, coffee shop, library, etc.) This also means that there is no direct oversight of online interactions removing some of the aversive control that may help to keep some students on task. This is to say, students may be watching TV, eating, talking with friends, or engaging in any number of competing activities while they run through an online lesson. The subjects in the postfeedback delay experiments previously cited did not have the opportunity to engage in these behaviors. Therefore, even though these competing activities may have an important impact on performance, the investigators did not have an opportunity to assess this impact in their experiments. Fourth, in their first experiment Crosbie & Kelly (1994) chose 10-seconds as the post-feedback delay duration and found that it was ineffective as a punisher. However, if a subject’s average response latency time to submit an answer following the presentation of a question is greater than 10-seconds then it is not clear why the delay would function as a punisher. In other words, in theory, subjects could move through the program more quickly (the presumed reinforcer) if they raced on every question. Logically, it follows that the delay could turn into a conditioned reinforcer due to it’s pairing with a faster pace through the program. In fact, Kelly & Crosbie (1997) appeared to admit as much when they noted that all of their subjects worked faster under the delay conditions. Fifth, in their study Crosbie & Kelly (1994) controlled for differences in problem difficulty by randomizing their conditions across sequential question sets. There is nothing inherently wrong with the appropriateness of this procedure. However, using problems of varying difficulty prevents the researchers from making any meaningful comparisons between time-based measures assessing how long it took the subjects to answer each question.

PURPOSE

Three experiments were conducted to assess the impact of postfeedback delays on racing during an online lesson and address some of the limitations inherent in the previous research on this topic. The research on post-feedback delays has been conducted only in highly controlled laboratory settings where subjects have no access to competing activities. However, today most computer-based learning that takes place online in a setting of the student's choosing. To more closely mimic the impact of post-feedback delays in these settings all three experiments were conducted online.

Additionally, with the exception of Crosbie & Kelly (1994), all of the research on postfeedback delays has confounded postfeedback delays with exposure to extra instructional content. To address this issue the following experiments isolated postfeedback delays to determine their impact alone on performance.

Another limitation addressed was the impact of relatively postfeedback delays investigated in previous studies (5 and 10 seconds). In contrast the following experiments increased the delay to 60 seconds for Experiments 1 and 2 and 30 seconds for Experiment 3. This duration was chosen for a few reasons. First, Crosbie & Kelly (1994) found 10-second postfeedback delays did not function as negative punishers following incorrect responding on a computer based lesson. Second, Zimmerman & Baydan (1963) reported that accuracy on a matching-to-sample task with college students improved as a function of longer time-out durations following incorrect responding until they reached 120-seconds, at which time the delay began negatively impacting the rate of correct performance. Similar conclusions supporting a moderate time-out duration of 10 to 60 seconds were reached by other researchers (Ferster & Appel, 1961; Zimmerman & Ferster, 1963) during investigations of variable time-out durations following inaccurate responding in a matching-to-sample task with pigeons. Based on these findings, it is reasonable to conclude that a 60-second delay duration should maximize accuracy without negatively impacting the rate of a subject's behavior.

Previous investigators researching postfeedback delays have relied solely on Holland & Skinner's programmed text, *Analysis of Behavior* for their instructional content. These materials mimic course content well but do not allow the experimenter to control for question complexity. In the following experiments 60 highly similar order of operations problems were utilized. This form of instructional content was chosen for several reasons. First, by using 60 similarly structured problems the level of question complexity can be controlled across the study. Furthermore, any potential practice effects can be counter-balanced by randomly altering the control, contingent delay, and contingent interactive delay conditions used in this study. Second, many common calculators are not easily designed to handle order of operation problems. Furthermore, by creating both experiments in Adobe Flash subjects are unable to copy the problem text and paste it into an online calculator. Third, data collected before running the experiments indicated that the average time to complete problems of this type was around 15-20 seconds. This is about the same amount of time it would take to type in and verify the accuracy of hand-entering the problem text in a calculator or online program. Finally, order of operation problems involve simple calculations but can still require a relatively significant response effort on behalf of subjects. This response effort can help show a clear distinction between racing versus actual possible attempts to correctly answer a problem. For example, depending on the problem response latencies of less than 5-seconds may be a fairly clear indication that racing took place, whereas, response latencies over 10 seconds may indicate at least some possibility that there may have been an attempt to correctly answer the problem.

EXPERIMENT 1

METHOD

Subjects

Twenty English-speaking undergraduate students over the age of 18 (17 female and 3 male) enrolled in a psychology course at the University of Nevada, Reno during the Fall 2011 semester served as subjects in this study. As compensation for participating, subjects received 1.5 research credits via the University's online human subject management system. The study took place online in a location of the subject's choosing. Before beginning, subjects were informed that they must complete the study in one sitting and finish it in its entirety to receive research credit. Subjects signed up for the study using the University's human subject pool management software program. After signing-up they were sent an email containing a link to a URL address where the online lesson was hosted. At the start of the study, subjects were given an additional opportunity to provide consent to volunteer before taking a tutorial designed to introduce them to the types of math problems and contingencies they were to encounter during the study.

Materials

This study was programmed using Adobe® Captivate® 5 e-learning authoring software. Adobe® Captivate® is a popular program used by instructors working in education and industry for creating online instructional content. The software is readily accessible to non-programmers while still allowing a high degree of flexibility for more sophisticated users. Online lessons are created using the software in a slide format similar to Microsoft's® PowerPoint® program. Unique to Adobe® Captivate is the ability to embed multiple question types (i.e., multiple choice, fill-in the blank, short answer, true/false, matching, Likert scale, hotspot, etc.) directly within a lesson, create branching scenarios, develop individualized feedback based on an end user's responses, and collect data on user interaction. Lessons created using the software are primarily published as .swf files uploaded to the internet and designed to play using Adobe's® ubiquitous Flashplayer software browser add-on. For the current experiment, the software was used to create an analog to an online lesson. This lesson was comprised of a short information and tutorial section followed by a series of 60 order of operation math problems. Data collected during the study with the widget was uploaded as a text file to a secure server for later downloading and analysis by the researchers.

Procedure

This study investigated the impact of control, contingent delay, and contingent interactive delay conditions on performance using an alternating-treatments design across 60 order of operations math problems.

Tutorial. Before encountering the 60 order of operations math problems, subjects were brought through a tutorial. During the tutorial, subjects were given three rules describing the consequences accompanying each condition when they answered a math problem incorrectly. The first rule stated that when the background screen color was white and a problem was answered incorrectly the program would simply move on to the next question. The second rule stated that when the background screen color was blue and a problem was answered incorrectly that the program would pause for 60 seconds. Finally, the third rule stated that when the background screen color was gray and a problem was answered incorrectly the program would pause for 60 seconds and every 5 seconds a button would need to be clicked to keep the 60 second countdown timer going. Subjects were also informed that when they answered a problem correctly they would move on to the next problem regardless of the background color of the screen. After being shown these rules, subjects were exposed to the consequences described in the rules when the background screen color was blue and gray. Once the tutorial was completed, subjects were presented with the 60 order of operations problems. Throughout the study, on all of the screens depicting an order of operations math problem a counter was shown in the bottom left corner of the screen (e.g., “16 out of 60”) notifying subjects how many more problems they had still to complete in the study.

Instructional Content. The 60 math problems used in this study were designed to assess a subject’s ability to apply the rules of order of operations. Each order of operation math problem used in the study was only 7 integers long, used the number 2 as an integer, resulted in no negative numbers or fractions during calculations, and contained multiple parentheses, addition, subtraction, multiplication, and division operations (e.g., 2x(2÷2)+2-2÷2+2). Finally, 10 problems each were generated that equaled, 1, 2, 3, 4, 5 & 6 for a total of 60 problems.

Independent Variables

The independent variables employed in this study were the rules and consequences for answering math problems incorrectly during the contingent delay and contingent interactive delay conditions. Each of these experimental conditions and a control condition were randomly assigned 20 of the 60 order of operation problems. Across all three conditions correct answers to a problem produced immediate feedback, no enforced delay, and the next problem. Incorrect answers to problems in the control condition produced the same consequence of immediate feedback followed by no enforced delay. Problems answered incorrectly in the contingent delay condition resulted in immediate feedback followed by an enforced 60-second delay. At the start of the delay the previously answered problem would disappear from the screen and be replaced by large text counting down 60-seconds. During the delay subjects could not progress in the lesson and had to wait until the 60-second timer finished counting down. Once the countdown was complete a “Click here to continue” button was shown on the screen. Subjects could progress to the next problem after clicking this button. Problems answered incorrectly during the contingent interactive delay condition were followed by a similar consequence. The only difference being that during the 60-second countdown the text timer would pause every 5-seconds and a randomly placed button labeled, “Click here to continue” would appear on the screen. When subjects clicked the button the countdown timer would continue until another 5-seconds had passed. When the countdown timer reached zero following this pattern and the subject pressed the “Click here to continue” button the next problem appeared on the screen.

In the program control condition problems were signaled by a white background color screen, contingent delay condition problems were signaled by a blue background color screen, and contingent interactive delay problems were signaled by a gray background color screen. It was hypothesized that problems presented on blue and gray backgrounds would be a stimuli in the presence of which racing would be punished, whereas the problems presented on a white background would be stimuli in the presence of which racing would not be punished. Under this preparation, the presumed reinforcer withheld following quick careless responding during the experimental conditions was progress through the lesson. The assumption was that some subjects would find the lesson content aversive (e.g., boring, difficult, etc.) and that by submitting answers as quickly as possible without concern for accuracy (i.e. racing) they could make fast progress through the lesson and escape the lesson content sooner. The delays were designed to eliminate that contingency by forcing the subject to stay in the instructional context when they answered a problem incorrectly. There was a concern that implementing a delay following incorrect responding would serve only to allow the subject an opportunity to access other competing activities (e.g., check email, watch TV, bathroom break, etc.). That is to say, even if the delay functioned as a time-out from access to the presumed reinforcer of progress through the lesson it would not serve as a time-out from other potential reinforcers in the subject's environment. The contingent interactive delay condition was designed to address this specific issue. It was hypothesized that by forcing the subject to interact every 5-seconds during the delay countdown this might limit their access to competing activities.

Dependent Measures

The dependent measures taken in this study included cumulative accuracy of responding and response latency measures. Cumulative accuracy measures allowed the researchers to track changes in accuracy of performance across the 3 conditions throughout the study. Additionally, response latency measures were collected for each problem and consisted of the time between the presentation of the problem on the screen until the time the subject pressed the "Submit" button to register their answer (Johnston & Pennypacker, 1993). Measuring response time in this way allowed the researchers to gain a rough estimate of the amount of time subjects allocated to solving problems in each condition and whether the subjects were racing (defined as submitting an answer within 5 seconds of question presentation).

RESULTS

Results from this study are graphically displayed as cumulative accuracy across conditions (e.g., Figures 1a, 2a, 3a, etc.) and response latency measures across conditions (e.g., Figures 1b, 2b, 3b, etc.). Cumulative accuracy rather than accuracy was chosen as the unit for the y-axis because intervention effects are easier to detect over the course of a study when subject's responses can only be categorized as correct or incorrect (Cooper, Heron & Heward, 2007). Visual analysis of cumulative records takes place by analyzing the slope connecting two data points. If no slope is observed (i.e., horizontal line) then no response was recorded. Additionally, the x-axis on both the cumulative accuracy and response latency measure graphs are grouped as trial sets rather than trails to correct for the distortion in slope that occurs when conditions are unevenly distributed across a study (e.g., when using an alternating treatments design). Each trial set located on the x-axis of these graphs is comprised of 1 trial from each condition. The first trial set displays a subject's performance on the first control, contingent delay and contingent interactive delay condition they encounter (i.e., Trials 1, 2 & 4). The second trial set displays a subject's performance on the second control, contingent, and contingent interactive delay condition they encounter (i.e., Trials 3, 5 & 8). This pattern is continued for twenty trial sets. Since conditions are randomized but are not separated by more than a few trials before the next instance, grouping conditions in this way allows for easier visual analysis of trends over the course of the study. Furthermore, although response latency measures do not require an analysis of slope to detect the existence of intervention effects these graphs still adopted the same x-axis so easy visual comparisons could be made with the cumulative accuracy graphs. For example, if a cumulative accuracy graph depicts a horizontal line across trial sets 13-18 in the control condition (i.e., all incorrect answers) a quick glance at the same trial sets on the response latency measure graph will quickly tell you whether the subject was scoring poorly because they were likely racing (i.e., recording response latencies at less than 5-seconds).

Subject Grouping

Results collected from the subjects revealed particular patterns of responding. Based on these patterns, subjects were grouped into 3 categories: Proficiency Group, Avoidance Group, and Endurance Group.

Proficiency Group. To be categorized in the Proficiency Group, subjects had to have scored 85% or better on all of the study questions (i.e., more than 50 out of 60 problems answered correctly) and have no more than 1 response latency measure recorded at less than 5 seconds (i.e., racing). Subjects categorized in this group did not race and performed very well on the order of operations problems. Of the twenty subjects included in this study, 14 were categorized in the Proficiency Group (Subjects 1, 2, 6- 8, 10- 12, 14, 15, 17-20). Twelve of these subjects were female and 2 were male. Visual analysis of the cumulative accuracy data for this group reveal relatively steady increasing trends across all conditions indicating consistently accurate performance throughout the study (see Figures 1a - 14a). Response latency measures recorded for subjects in this group indicate relatively stable response times for Subjects 1, 2, 7, 10, 11, 15, 17, 19 & 20 (see Figures 1b, 2b, 4b, 6b, 7b, 10b, 11b, 13b & 14b) occasionally following some initial variability on the first few problems for Subjects 1, 7 & 19 (see Figures 1b, 4b & 13b). Response latency measures for the other subjects in this group spanned a wide range. Subject 6's response latency measures (see Figure 3b) displayed a high variable level of responding with zero trend. Subject 8's response latency measures (see Figure 5b) displayed a very high variable level of responding with a slight decrease in trend during the last third of the study. Subject 12's response latency measures (see Figure 8b) displayed high variability with a decreasing trend throughout the study. Subject 14's response latency measures (see Figure 9b) displayed high variability with a slight decrease in trend over the course of the study. Finally, Subject 18's response latency measures (see Figure 12b) displayed a very high variable level of responding with a decrease in trend over the course of the study.

Avoidance Group. To be categorized in the Avoidance Group, subjects had to have scored less than 85% on all the study questions and have more than 1 response latency measure recorded at less than 5-seconds. Subjects categorized in this group were prone to racing during the study and were therefore not as accurate as subjects in the Proficiency group. Of the twenty subjects included in this study, 5 were categorized in the Avoidance Group (Subjects 3-5, 9 & 13). Four of these subjects were female and 1 subject was male. Cumulative accuracy data for Subject 3 shows a decrease in performance during the control condition about 1/3 of the way through the study (see Figure 16a). This same decrement in performance was not seen during time-out and interactive time-out conditions for this subject, with performance during these conditions remaining stable and increasing at a steady rate throughout the study. Response latency data for Subject 3 confirm that performance worsened in the control condition when response latency measures were recorded at less than 5-seconds (see Figure 16b). Response latency data during the contingent delay and contingent interactive delay conditions for this subject reveal a moderate to high variable level of responding with zero trend throughout the study.

Cumulative accuracy data for Subject 4 shows a relatively accurate steady performance for two thirds of the study across all conditions (see Figure 17a). However, with about 1/3 of the study left, a decrement in accuracy is seen during the control condition while accuracy during the two experimental conditions remains high. Response latency measures confirm that response times were less than 5-seconds during this time in the control condition with the exception of the last two control condition trials where the data indicate that the subject began spending a bit more time before answering the problem (see Figure 17b). Response latency measures taken during the experimental conditions indicate moderate variability with a very slight decreasing trend over the course of the study.

Cumulative accuracy and response latency data for Subject 5 showed a clear distinction between performance in the control condition and the two experimental conditions (see Figure 18a). Specifically, the data for this subject indicate poor performance within the control condition and relatively accurate performance in the two experimental conditions throughout the study. Furthermore, a quick glance at the response latency measures recorded for this subject indicate that on all but the first two control condition trials they responded to the question in less than 5-seconds (see Figure 18b).

Cumulative accuracy data for Subject 9 shows relatively poor performance across conditions, especially during 6 trials in the middle of the study during the contingent delay condition (see Figure 19a).

An analysis of the response latency measures recorded during this poor run shows that with the exception of two trials, the subject's response latency measures were equivalent with their response times during the other conditions occurring about the same time (see Figure 19b). This indicates that rather than guessing on contingent delay trials during this middle run the subject was likely trying to answer the question correctly but failed to do so. Response latency measures for this same subject indicate that towards the very end of the study, specifically the last three trials in the control condition and the last trial in the contingent delay condition, the subject was responding to questions within 5 seconds indicating racing took place. This same quick responding was not present at the end of the study during the contingent interactive delay conditions (but did appear once in trial set 15) demonstrating one of the only times there was notable differentiation between performance in the contingent delay and contingent interactive delay conditions.

Cumulative accuracy data for Subject 13 indicate very poor accuracy during the study (see Figure 20a). Response latency measures are very highly variable across all conditions (see Figure 20b). Additionally, periodically throughout the study the subject's response times were less than 5-seconds in the control condition. During the final four trial sets response latency measures for control and contingent delay conditions were less than 5-seconds indicating the only other time differentiation was observed between the two experimental conditions.

Endurance Group. To be categorized in the Endurance Group, subjects had to have scored 85% on all the study questions and have no more than 1 response latency recorded at less than 5-seconds. Subjects in this group did not race but still performed poorly on the order of operations problems. Only 1 male subject out of the 20 subjects included in this study met the criteria necessary to be included in the Endurance Group. Subject 16's cumulative accuracy of responding (see Figure 22a) was relatively low across all three conditions yet this subject only showed a moderate level of variability in their response time and only once, very early on in the study, did this subject's response time drop below 5-seconds (see Figure 22b).

DISCUSSION

Results collected from the experiment indicate that the intervention was effective in preventing racing for five of the subjects in the Avoidance Group. More notably is the 14 subjects in the Proficiency Goup who did not appear to need the intervention to perform at optimal or near optimal levels. With such a high level of performance occurring across so many subjects without the need for the intervention it was believed that a ceiling effect was encountered. One way to address this issue is to increase the response effort involved in solving each problem by making each question more difficult. By requiring subjects to spend more time solving each problem racing was hypothesized to be more likely to occur. To address this question a similar second experiment was conducted with more difficult order of operations problems.

EXPERIMENT 2

METHOD

Subjects

Twenty subjects were recruited to participate in this study (15 female and 5 male). All subjects met the same criteria and were compensated in the same way as described in the first study.

Materials

This study was also programmed using Adobe® Captivate® 5 e-learning authoring software.

Procedure

During this study subjects were exposed to the same procedure as described in the first study. The only difference was the modification of the 60 order of operations problems used in the study. In the previous study, each problem contained 7 integers (e.g., 2+2x(2-2÷2)÷2x2). In the current study this number was changed to 10 integers (2x(2+2)÷2-2+(2x2÷2)-2+2) in an attempt to raise the difficulty and thus the response effort required to answer each problem. Additionally, to further control the difficulty of each problem the number of signs was more systematically applied. Specifically, each order of operation math problem used in the study was only 10 integers long, used the number 2 as an integer, resulted in no negative numbers or fractions during calculations, and contained two sets of parentheses, two multiplication and division operations, and a mix of five addition and subtraction operations (e.g., (2-2)÷2+2x2-2x(2-2+2)÷2). Finally, 10 problems each were generated that equaled, 1, 2, 3, 4, 5 & 6 for a total of 60 problems.

Independent Variables

The same independent variables described in the first study were used here. Even the randomized order of conditions within the alternating treatments design remained the same.

Dependent Measures

Dependent measures collected for this study were also the same (i.e., cumulative accuracy, response latency measures) as the first study.

RESULTS

Results for Experiment 2 are graphically presented using the same format described earlier for Experiment 1.

Subject Grouping

Subjects in this experiment were grouped using the same 3 categories described in Experiment 1 (i.e., Proficiency Group, Avoidance Group, and Endurance Group) with the addition of one new category. The Inattentive Group categorization was added to account for the performance of subjects whose accuracy across the study was less than 50% (i.e., less than 30 out of 60 problems correct) and whose data set included 5 or more response latencies measures under 5-seconds across all 3 conditions. Subjects in the Inattentive Group were characterized by performance that showed guessing across all conditions throughout the experiment.

Proficiency Group. Of the 20 subjects included in Experiment 2, eight female subjects were categorized in the Proficiency Group (Subjects 23, 26, 27, 29, 30, 32, 36 & 39). Visual analysis of the cumulative accuracy data for this group reveal steady increasing trends across all conditions indicative of consistently accurate performance throughout the study (see Figures 24a - 31a). Response latency measures recorded for subjects in this group show relatively stable response times for Subjects 26, 29, 30 & 32 (see Figures 25b, 27b, 28b, 29b) and a wider range of variability in response times for Subjects 23, 27, 36 & 39 (see Figures 24b, 26b, 30b & 31b). Response times for Subjects 23, 30, 32 & 39 (see Figures 24b, 28b, 29b & 31b) showed a slight decrease in trend over the course of the study. A similar, but more moderate, trend appears for Subjects 26, 27, 29 & 36 (see Figures 25b, 26b, 27b & 30b) indicating faster performance the longer these subjects were in the study (i.e. practice effects).

Avoidance Group. Of the twenty subjects included in this study, 8 subjects (6 female and 2 male) were categorized in the Avoidance Group (Subjects 21, 22, 28, 31, 33, 34, 37 & 40). Cumulative accuracy data for Subject 21 indicate that less than halfway through the study they incorrectly answered every problem in the control condition while accuracy in the two experimental conditions remained high and stable (see Figure 33a). A glance at the response latency data for this subject shows that with the exception of trial set 12, accuracy dropped during control trials when response times were under or around 5-seconds (see Figure 33b).

Cumulative accuracy data for Subject 22 shows a similar pattern, with inaccurate, unstable performance occurring during control condition trials around 1/3 of the way through the study (see Figure 34a). Inaccurate performance maps on directly to response latency measures around 5-seconds that occurred sporadically during the last two-thirds of the study for control condition problems (see Figure 34b). Performance in the control condition for Subject 22 can be directly compared to their performance in the two experimental conditions, which indicated high accurate stable performance with a steady decreasing trend in response times over the course of the study. Additionally, no response times for Subject 22 in either experimental condition was less than or even around 5-seconds.

Cumulative accuracy data for Subject 28 shows a very clear differentiation between their performance in the control condition and the two experimental conditions (see Figure 35a). Throughout the study answers submitted during the control condition trials were inaccurate while answers submitted during the two experimental conditions were highly stable and accurate. Additionally, Subject 28's response latency data were much lower in the control conditions hovering around 5-seconds or under for the vast majority of trials during the study while response times in the two experimental conditions were steady with a slight decreasing trend hovering around 30-40 seconds for each problem (see Figure 35b).

Cumulative accuracy for Subject 31 show low accuracy across conditions (see Figure 36a). However, response latency data for this subject indicate that about halfway through the study they began responding around 5-seconds during control condition problems only (see Figure 36b). Response latency measures for the two experimental conditions remained moderately variable (more so across the contingent delay condition) with a very slight decrease in trend.

Cumulative accuracy data for Subject 33 shows a highly accurate performance across all conditions for almost half of the study (see Figure 37a). However, around the halfway point the subject's performance became very inaccurate during control condition problems while accuracy in the experimental conditions remained high. Inaccuracy during control condition problems can be traced directly to the subject's response latency data, which hovers around 5-seconds for the second half of the study and remains stable with zero trend above 20-seconds in the two experimental conditions (see Figure 37b).

Cumulative accuracy data for Subject 34 shows moderate performance across all conditions with accuracy becoming worse during control condition problems halfway through the study (see Figure 38a). Response latency data for this subject show significantly lower response times during control condition problems for this subject in the second half of the study (see Figure 38b). Additionally, this subject appears to have raced on two contingent delay problems over the course of the study.

Cumulative accuracy data for Subject 37 show good performance in the study across all three conditions (see Figure 39a). Subject 37 is an outlier from the rest of the group in that the majority of response latency times under or around 5-seconds occurred during the first 3 contingent delay trials (see Figure 39b). Following this rough beginning the subject's response times were highly stable with a slight decreasing trend around 20 seconds.

Cumulative accuracy data for Subject 40 shows a very clear differentiation between their performance in the control condition and the two experimental conditions (see Figure 40a). Throughout the study answers submitted during the control condition trials were inaccurate while answers submitted during the two experimental conditions were highly stable and accurate. Additionally, Subject 40's response latency data were much lower in the control conditions hovering around 5-seconds or under for the vast majority of trials during the study while response times in the two experimental conditions were steady with a slight decreasing trend hovering around 35-40 seconds for each problem (see Figure 40b).

Endurance Group. Only 3 (2 female and 1 male) out of the 20 subjects included in Experiment 2 met the criteria necessary to be included in the Endurance Group (Subject 24, 25 & 35). All three subjects performed with moderate accuracy across all conditions (see Figures 42a, 43a & 44a). Response latency data for Subjects 24 & 25 was highly stable and hovered around 15-25 seconds (see Figures 42b & 43b). Subject 35's response latency data was highly variable at the start of the study across conditions but stabilized about halfway through (see Figure 44b). None of these subjects had more than one response latency less than 5-seconds.

Inattentive Group. Only one male subject (Subject 38) met the criteria necessary to be included in the Inattentive Group. This subject's accuracy was very poor across all conditions throughout the study (see Figure 46a). Response latency data for Subject 38 was highly variable across conditions (see Figure 46b). About halfway through the study response latency data during control condition problems dropped below or around 5-seconds. About 2/3 of the way through the study response latency measures in the two experimental conditions followed suit and also dropped below or around 5-seconds.

DISCUSSION

Results collected from the second experiment show that the intervention was effective in preventing racing for eight of the subjects in the Avoidance Group. Additionally, the number of subjects included in the Proficiency Group dropped by almost half from the first experiment to the second. These results indicate that increasing the number of integers per problem had the intended effect of making each problem more difficult. Although the results from both experiments confirm that contingent postfeedback delays can prevent racing during an online lesson there are some limitations that bear mentioning. First, each contingent delay period was lengthy and could increase time in the instructional context substantially for low performers. Second, no attempt was made to determine whether the contingent interactive delay actually prevented subjects from contacting outside competing activities. For example, when encountering an contingent interactive delay a subject could have engaged in any number of competing activities as long as they quickly interacted with the screen at some point after 5-seconds had passed to advance the countdown timer. With no way of determining whether the contingent interactive delay condition prevented subjects from accessing competing activities there is no way to assess whether it is even important to do so to prevent racing. Third, no social validity data were collected to assess subjects' preferences for the experimental conditions. Since these conditions involve using aversive delays to control performance, it is unlikely that subjects would have rated their use highly. However, it would have been interesting to see just how aversive these subjects considered these conditions especially when compared to one another. To address these limitations a third experiment was conducted.

EXPERIMENT 3

METHOD

Subjects

Thirty subjects were recruited to participate in this study (18 female and 12 male). All subjects met the same criteria and were compensated in the same way as described in the two previous experiments.

Materials

Like the two previous experiments, this study was also programmed using Adobe® Captivate® 5 e-learning authoring software.

Procedure

During this study subjects were exposed to a similar procedure as described in the second experiment with some important exceptions. First, multiplication signs used in the order of operation problems were converted from an “x” sign into a “·“ (i.e., interpunct) sign . This change was designed to address concerns that the “x” sign implies the subject must solve for a variable. Second, the contingent interactive delay condition no longer required subjects to press a randomly placed button every 5-seconds to advance the countdown timer. Instead, during the 60-second countdown a button label shown on the screen changed from “Look for the button to appear here” to “Click here to go to the next problem” at a random time somewhere within 20-40 seconds into the delay countdown for only 3-seconds (30 seconds averaged across the study). If the subject pressed this button when the label changed they were moved onto the next problem. However, if they did not press the button before it disappeared they were forced to redo the same delay from the beginning. Once the countdown finished, subjects were given access to another button labeled, “Click here to restart the countdown” that when clicked started the delay over again. Third, instead of using a 60-second pause for the contingent delay condition used a 30-second delay. Fourth, at the end of the experiment subjects were asked 3 social validity questions. These questions asked subjects to use a Likert scale to rate how satisfied they were with the consequences for incorrect answers in each of the three conditions (e.g., “On a scale of 1-9, how satisfied were you with the consequence for incorrect answers when the screen was BLUE?”). After answering these questions subjects were asked to self-report what they did when, or if, they encountered a delay or interactive delay during the experiment (e.g., “If you answered a problem incorrect when the screen was GRAY what did you do while you waited for the pause to end?”). After answering these questions, subjects were exposed to a forced choice condition where they were given the opportunity to choose to complete one last problem in the experiment in a contingent delay or an interactive contingent delay condition. Finally, subjects were then given 2 blank text boxes to provide self-reports to the following questions (1) "How did you solve the math problems in the study (e.g., in head, on paper, using a calculator, etc.)?" and (2) "Where did you complete this study?"

Independent Variables

The same independent variables described in the two previous experiments were used here. However, each condition was randomly re-assigned to 20 order of operations problems each and the 60 problems were randomized within the alternating treatments design.

Dependent Measures

In addition to the new self-report measures mentioned in the procedure, dependent measures collected for this study remained the same (i.e., cumulative accuracy, response latency measures) as the two previous studies.

RESULTS

Results for Experiment 3 are graphically presented using the same format described earlier for the two previous experiments.

Subject Grouping

Subjects in this experiment were grouped using the same 4 categories described in Experiment 2 (i.e., Proficiency Group, Avoidance Group, Endurance Group, and Inattentive Group).

Proficiency Group. Of the 30 subjects included in Experiment 3, eleven subjects (6 females and 5 males) were categorized in the Proficiency Group (Subjects 41, 42, 46, 47, 49, 51, 52, 56, 60, 63 & 66). Visual analysis of the cumulative accuracy data for this group reveal steady increasing trends across all conditions indicative of consistently accurate performance throughout the study (see Figures 48a, 49a, 50a, 51a, 52a, 53a, 54a, 55a, 56a, 57a & 58a). Response latency measures recorded for subjects in this group show relatively stable level response times for Subjects 41, 42, 46, 47, 51, 52, 56 & 66 (see Figures 48b, 49b, 50b, 51b, 53b, 54b, 55b, 58b) and a wider range of variability in response times for Subjects 49, 60 & 63 (see Figures 52b, 56b & 57b). Only Subject 63's performance indicates a practice effect as response times generally appeared to get shorter over the course of the study (see Figure 57b).

On the self-report question (see Figure 84a), "On a scale of 1-9, how satisfied were you with the consequence for incorrect answers when the screen was [WHITE/BLUE/GRAY]?" Subjects 41, 42, 46, 51, 60, and 63 all reported a higher rating for the consequences on a white screen (i.e., control condition problems) than the blue and gray screens (i.e., contingent delay & contingent interactive delay conditions). Subject 49 reported the lowest rating for all 3 conditions and Subject 52 reported the highest rating for all 3 conditions. Subject 56 gave the interactive contingent delay condition the highest rating of the three options and Subject 63 did the same with the contingent delay condition. For the Proficiency Group only 3 out of 6 of the subjects (Subjects 49, 60 & 66) choose to complete an extra contingent interactive delay problem over a contingent delay problem during the forced choice condition.

On the self-report question (see Table 1), "If you answered a problem incorrect when the screen was BLUE what did you do while you waited for the pause to end?" 4 of the subjects reported engaging in off-task behavior (Subjects 46, 47, 56 & 60) during the delay, 5 subjects reported on task behavior (Subjects 41, 42, 49, 51 & 63), and for 2 subjects it was unknown whether they were reporting on-task or off-task behavior (Subjects 52 & 66). In response to this same question directed at the GRAY screen (i.e., the contingent interactive delay condition), only Subject 60 reported off-task behavior during the delay, 8 subjects reported on-task behavior (Subjects 41, 42, 46, 49, 51, 56, 63 & 66), and for 2 subjects it was unknown whether they were reporting on-task or off-task behavior.

On the self-report question (see Table 2), "How did you solve the math problems in the study (e.g., in head, on paper, using a calculator, etc.)?" over half of the subjects reported using a calculator at least some of the time (Subjects 41, 47, 52, 56, 60 & 66) while the rest of the subjects reported completing the problems in their head and/or on paper (Subjects 42, 46, 49, 51 & 63).

On the self-report question (see Table 3), "Where did you complete this study", subjects in the Proficiency Group reported completing the study in their bedrooms (Subjects 46, 47, 49, 51, 60, 63 & 66) or the library (Subjects 41, 42, 52 & 56).

Avoidance Group. Seven subjects (5 females & 2 males) out of 30 met the criteria to be included in the Avoidance Group in Experiment 3 (Subjects 43, 53, 55, 58, 59, 69 & 70). Cumulative accuracy data for Subject 43 indicate poor performance across all 3 conditions throughout the study (see Figure 60a). Visual analysis of this subjects response times show that about 1/3 of the way through the study the subject began racing during only control condition problems (with the exception of trial set 10) while maintaining a longer response latencies across the two experimental conditions (see Figure 60b).

Cumulative accuracy data for Subject 53 indicate near perfect performance on contingent interactive delay problems and poor performance on both contingent delay and control condition problems during the study (see Figure 61a). Response latency measures for this subject show highly variable responding at the beginning of the study across all three conditions. As the study progressed a decreasing trend can be seen with performance stabilizing. During the last third of the study Subject 53 shows racing behavior on almost control condition problem they encountered (see Figure 61b).

Cumulative accuracy data for Subject 55 indicate relatively poor performance across all three conditions, especially the control condition problems (see Figure 62a). Response latency data for this subject show racing behavior on almost every control condition problem with moderately decreasing variable response times for the two experimental conditions throughout the study (see Figure 62b).

Cumulative accuracy data for Subjects 58 & 59 indicate moderate to poor performances across all three conditions, especially the control condition problems during the last half of the study (see Figures 63a & 64a). Response latency measures for these subjects show racing occurring sporadically across all three conditions during the first half of the study, while during the second half racing occurs consistently only during control condition problems (see Figure 63b & 64b).

Cumulative accuracy data for Subject 69 indicate moderate performance across the two experimental conditions and extremely poor performance on control condition problems (see Figure 65a). Visual analysis of response latency data for this subject show highly variable response times for the two experimental conditions and racing during almost every control condition problem (see Figure 65b).

Cumulative accuracy data for Subject 70 indicate moderate performance across the two experimental conditions but poor performance on control condition problems for the first two-thirds of the study (see Figure 66a). Response latency data for this subject show variable levels of responding with a slight decrease in trend across the two experimental conditions. Additionally, for the first two-thirds of the study racing occurred on almost every control condition problem (see Figure 66b).

On the self-report question (see Figure 84a), "On a scale of 1-9, how satisfied were you with the consequence for incorrect answers when the screen was [WHITE/BLUE/GRAY]?" 5 of the 7 subjects in the Avoidance Group (Subjects 43, 53, 58, 59 & 69) reported a higher rating for the control condition consequences, followed by the contingent delay consequences, and finally the contingent interactive delay problems. Subject 55 reported low ratings across all three conditions but in the opposite order. Subject 70 gave the consequences for all three conditions the lowest rating available. For the Avoidance Group only 2 out of 7 of the subjects (Subjects 43 & 59) choose to complete an extra contingent interactive delay problem over a contingent delay problem during the forced choice condition.

On the self-report question (see Table 1), "If you answered a problem incorrect when the screen was BLUE what did you do while you waited for the pause to end?" 6 of the 7 subjects in the Avoidance group

(Subjects 43, 53, 55, 58, 59 & 70) reported off-task behavior and 1 subject (Subject 69) reported on-task behavior during the contingent delay condition. In response to this same question directed at the GRAY screen (i.e., the contingent interactive delay condition) only Subject 53 reported off-task behavior while all of the other subjects in the Avoidance Group (Subjects 43, 55, 58, 59, 69 & 70) reported on-task behavior during the contingent interactive delay condition.

On the self-report question (see Table 2), "How did you solve the math problems in the study (e.g., in head, on paper, using a calculator, etc.)?" only Subject 43 reported using a calculator. All other subjects (Subjects 53, 55, 58, 59, 69 & 70) reported that they attempted to solve the problems in their head.

On the self-report question (see Table 3), "Where did you complete this study", subjects in the Avoidance Group reported completing the study in their bedrooms (Subjects 43, 55 & 70), on campus (Subject 53), in their dorm room (Subject 58), in the dining room (Subject 69), or the library (Subject 59).

Endurance Group. Six subjects out of the 30 met the criteria to be included in the Endurance Group (Subjects 44, 57, 61, 62, 65 & 68). Visual analysis of the cumulative accuracy data for subjects in this group reveal relatively low performance across all three conditions, especially for Subjects 44, 65 & 68 (see Figures 68a, 69a, 70a, 71a, 72a & 73a). Response latency data for subjects in the Endurance Group show moderate to highly variable levels of responding throughout the study across all conditions with only one instance of racing occurring for Subject 44 during a contingent interactive delay condition problem (see Figures 68b, 69b, 70b, 71b, 72b & 73b).

On the self-report question (see Figure 84c), "On a scale of 1-9, how satisfied were you with the consequence for incorrect answers when the screen was [WHITE/BLUE/GRAY]?" only Subjects 57 & 61 rated the control condition consequences higher than the two experimental conditions. Subjects 44, 62 & 68 rated all conditions basically equivalent to one another and Subject 65 only provided a low rating of 2 for contingent delay consequence. For the Endurance Group only 2 out of 6 of the subjects (Subjects 44 & 61) choose to complete an extra contingent interactive delay problem over a contingent delay problem during the forced choice condition.

On the self-report question (see Table 1), “If you answered a problem incorrect when the screen was BLUE what did you do while you waited for the pause to end?” 2 of the subjects reported off-task behavior (Subjects 44 & 57), 3 subjects reported on task behavior (Subjects 62, 65 & 68), and for Subject 61 it was unknown whether they were reporting on-task or off-task behavior. In response to this same question directed at the GRAY screen (i.e., the contingent interactive delay condition), all of the Subjects in the Endurance Group reported on-task behavior during the delay.

On the self-report question (see Table 2), “How did you solve the math problems in the study (e.g., in head, on paper, using a calculator, etc.)?” none of the subjects reported using a calculator and all of the subjects reported completing problems in their head sometimes with the assistance of paper.

On the self-report question (see Table 3), “Where did you complete this study”, subjects in the Endurance Group reported completing the study in the library (Subjects 44, 57 & 68) their bedroom (Subject 65), at work (Subject 62), or in their dorm room (Subject 61).

Inattentive Group. Six subjects out of the 30 met the criteria to be included in the Inattentive Group (Subjects 45, 48, 50, 54, 64 & 67). Visual analysis of the cumulative accuracy data for all subjects in this group reveal extremely low performance across all three conditions (see Figures 75a, 76a, 77a, 78a, 79a & 80a). Response latency data for subjects in the Inattentive Group show numerous racing attempts across all three conditions (see Figures 75b, 76b, 77b, 78b, 79b & 80b). Only Subject 54 showed some variation during the course of the study where racing did not occur.

On the self-report question (see Figure 84d), “On a scale of 1-9, how satisfied were you with the consequence for incorrect answers when the screen was [WHITE/BLUE/GRAY]?” every subject in the Inattentive Group rated the control condition consequence higher than the consequences affiliated with the two experimental conditions. The only exception was Subject 54 who rated each condition equally at 5. Four of the 6 subjects rated the contingent delay condition higher than the contingent interactive delay condition (Subjects 45, 48, 50 & 64). Only Subject 67 rated the contingent interactive delay condition higher. For the Inattentive Group only 2 out of 6 of the subjects (Subjects 54 & 64) choose to complete an extra contingent interactive delay problem over a contingent delay problem during the forced choice condition.

On the self-report question (see Table 1), “If you answered a problem incorrect when the screen was BLUE what did you do while you waited for the pause to end?” 5 of the subjects reported off-task behavior (Subjects 45, 48, 50, 64 & 67) and 1 subject (Subject 54) reported on-task behavior during the delay. In response to this same question directed at the GRAY screen (i.e., the contingent interactive delay condition), all of the Subjects in the Inattentive Group reported on-task behavior during the delay.

On the self-report question (see Table 2), “How did you solve the math problems in the study (e.g., in head, on paper, using a calculator, etc.)?” none of the subjects reported using a calculator and all of the subjects reported completing problems in their head sometimes with the assistance of paper.

On the self-report question (see Table 3), “Where did you complete this study”, subjects in the Endurance Group reported completing the study in their bedroom (Subject 45, 50 & 67), at home (Subject 48), in their living room (Subject 64) or in their dorm room (Subject 61).

GENERAL DISCUSSION

Based on the results of these experiments three main conclusions can be drawn. First, enforced contingent delays can prevent racing during an online lesson for some students. Second, contrary to previous findings (Crosbie & Kelly, 1994; Johnson & Dickinson, in press; Kelly & Crosbie, 1997), the delays used to prevent racing do not need to include exposure to extra instructional content to be effective. Finally, limiting access to competing activities during enforced delays does not appear to impact racing behavior. Instead, simply withholding progress in the lesson alone is enough to deter racing behavior.

The results from all three experiments provide some evidence that contingent postfeedback delays following incorrect answers in an online assessment can reduce racing behavior. These results do not appear to be universal as subjects in the Inattentive Group for Experiments 2 (N = 1) and Experiment 3 (N = 6) were insensitive to the contingency as racing behavior took place across the control and experimental conditions. Additionally, although a direct comparison was never made, the effects of the intervention were not as apparent when the delay was reduced from 60 seconds to 30 seconds.

In both experiments racing was defined as a response latency measure of less than 5-seconds. Racing can be considered a form of guessing as it occurs when an answer is given before enough time has passed for an accurate answer to have been formulated. However, it is possible that subjects also guessed some of their answers after viewing a problem for a significant amount of time (e.g., after 10-seconds had passed). Although this is impossible to verify with the current measurement system, it seems unlikely that guessing occurred very often after a significant amount of time had passed in the current studies. In all three experiments subjects were exposed to problems that were highly similar throughout. This level of similarity should have eliminated any time that might otherwise had been used to scan a novel question and determine whether it was too difficult or time-consuming to answer. Very early on, subjects should have learned about how long each problem took to answer accurately and the degree of difficulty they would encounter. Without the need to assess the response effort required to answer each problem, when guessing did occur it more than likely happened within the first 5 seconds of encountering the problem. The significant discrepancy between trials assumed to be guessed (i.e., answered within 5 seconds of question presentation) and the response latency measures for other problems, where it is assumed subjects attempted to answer the question correctly, provide further evidence to support this assumption. Additionally, analysis of the accuracy of responding during trials where racing took place show correct responding on these questions at chance levels. The data also indicate that when subjects did answer incorrectly this was not always due to guessing (unless their answer was provided within the first 5 seconds). In other words, when subjects answered a problem incorrectly their response latency measures were either 0-5 seconds (i.e. racing) or at a level comparable to the response latency measures of when they answered a problem correctly (indicating a failed attempt). When subjects did race they overwhelming did so in the presence of control condition problems (see Figure 82a, 82b & 82c).

Proficiency Group. Although no significant differentiation in performance can be seen across conditions for subjects in the Proficiency Group in Experiment 1 (N=14), Experiment 2 (N=8), and Experiment 3 (N = 11) it would be inaccurate to draw the conclusion that contingent delays and contingent interactive delays were ineffective for these subjects. Instead, the subjects in this group demonstrate that the contingent delay and contingent interactive delay periods were not needed to maintain an optimal or near-optimal level of performance during the study (see Figures 15a, 15b, 32a, 32b, 59a & 59b). This is to say, the data show that it is unlikely that the delay periods were the variables responsible for controlling the subject's behavior of responding accurately in the presence of the study's problems. Instead, for these subjects their behavior of answering each problem correctly is likely the result of a long history of responding to questions demanding significant response effort. It is important to note that the number of subjects that were included in the Proficiency Group was reduced substantially when the problems in the second and third experiment were extended to accommodate 3 more integers. This change in problem complexity was also accompanied by a significant increase in the number of subjects that were categorized in the Avoidance Group. This indicates that although contingencies may be in place to maintain optimal performance for some subjects there is a breaking point where these contingencies could benefit from supplements like contingent delays.

Avoidance Group. The results from subjects categorized in the Avoidance Group in Experiment 1 (N=5), Experiment 2 (N=8), Experiment 3 (N = 7) provide evidence that the contingent delays, contingent interactive delays, and their associated rules can be used to increase accuracy and the amount of response time spent on a problem in an online lesson. Visual analysis of the cumulative accuracy and response latency measures of the performances of subjects in this group in all three experiments support this claim in a few ways. First, with the exception of Subject 9 in Experiment 1 and Subject 37 in Experiment 2 accuracy was greater in the experimental conditions when compared to performance on problems in the control condition (see Figures 21a, 41a & 67a). Second, when performance was inaccurate, an analysis of the response time for that same trial usually indicated the subject had submitted their answer in less then 5-seconds (i.e., racing). Third, in both experiments, with the exception of Subject 9 and Subject 37, the average response latency for all subjects during control condition problems was less than the the average response latency during experimental condition problems (see Figures 21b, 41b & 67b).

Endurance Group. Results from Endurance Group subjects across all 3 experiments (Subjects 16, 24, 25, 35, 44, 57, 61, 62, 65 & 68) emphasize that the contingent delays used in this study are not designed to teach pre-requisite skills needed to produce correct answers. It possible that contingent delays used over a longer period (e.g., throughout a semester) may result in students lacking pre-requisite skills spending more time studying or seeking out assistance throughout the semester in an effort to avoid the contingency during their assignments. However, the current study doesn't address this question. In the current experiments, subjects in this group appeared to lack the pre-requisite skills necessary to accurately answer the math problems in the study regardless of the condition (see Figures 23a, 45a, 74a). Even in the face of continual failure, these subjects did not record more than one response latency measure less than 5 seconds, although average response latency measures for this group (see Figures 23b, 45b & 74b) were relatively low when compared against data in Proficiency Groups (see Figures 15b, 32b & 59b). Response latency measures of performance for subjects in this group indicate that if racing was occurring it rarely happened within the first few seconds of encountering the problem, which likely indicates that participants were attempting to accurately answer every question they encountered even when a significant number of their answers were wrong. The relative stability of these measures throughout the study also indicates that continual exposure to the time-out and interactive time-out interventions did not punish their behavior of attempting to answer each question with the same amount of effort as for earlier problems. This is an important finding as it provides some evidence that the interventions may not be powerful enough to lead to learned helplessness (Hiroto & Seligman, 1975) in individuals who encounter it frequently.

Inattentive Group. During the first two experiments only 1 subject's (Subject 38) performance met the criteria to qualify for categorization in the Inattentive Group. This number was increased to 6 subjects out of 30 when the contingent delay and contingent interactive delay time period was reduced by half. Subjects in the Inattentive Group had extremely poor accuracy and multiple response latencies less than 5-seconds across all 3 conditions indicating that the intervention was not nearly as effective when comparing their data against the data of subjects in the Avoidance Group (see Figures 47a, 47b, 81a & 81b).

Learning vs. Attending. The contingent delay and contingent interactive delay interventions used in these studies were not designed to help with learning. Instead these interventions were designed to increase attending on the part of subjects who already had the necessary repertoire to respond accurately to questions in an online format but failed to apply those skills, assumedly to escape an aversive instructional context as quickly as possible. These subjects had the pre-requisite skills to do well, but did not apply themselves fully because they were trying to finish the lesson quickly. For these types of students it is assumed the instructional content is aversive in some way (e.g. difficult, time-consuming, or boring, etc.) and by racing or putting in less effort than is required to ensure a correct answer, they can escape the instructional context quickly. Enforcing a delay following incorrect responding disrupts that contingency by temporarily halting progress in the lesson (i.e., the presumed reinforcer) when racing occurs.

Although teaching subjects how to answer assessment questions correctly was not the purpose of the research presented here, it is easy to see how contingent postfeedback delays could be combined with other interventions to teach students new skills. In fact, until now, exposure to extra instructional content during contingent postfeedback delays was the reason why these interventions were considered effective at all (Crosbie & Kelly, 1994; Johnson & Dickinson, in press; Kelly & Crosbie, 1997). In other words, including postfeedback delays can increase attending and using those delays to create further opportunities for learning can assist with training new skills. For example, arranging contingencies such that a mastery level must be met before students can move on in a lesson could result in student's attending more closely to materials in order to complete a lesson faster. Additionally, requiring students to complete extra practice questions during a delay could also result in students racing less and attending more carefully in order to avoid having to complete more problems.

Limiting Access to Competing Activities. Data collected from all 3 experiments indicate little differentiation between the contingent delay and contingent interactive delay conditions. This finding indicates that using an enforced delay under the arranged conditions was sufficient to prevent racing without the need to require further interaction. This same finding was replicated even when a shorter delay was used in Experiment 3. The purpose of including a separate contingent interactive delay condition was to create a contingency where subjects would be forced to interact with the delay condition and therefore be limited from contacting reinforcers in their immediate setting. Cooper, Heron & Heward (2007) define time-out as "the contingent withdrawal of the opportunity to earn positive reinforcement or the loss of access to positive reinforcers for a specified time". For a procedure to truly qualify as time-out, there must be a clear discrepancy between the time-in and time-out conditions. Specifically, during the response-contingent time-out conditions reinforcement should be withheld and as a result the behavior of interest should decrease in frequency in the future. In the first two experiments, it was hypothesized that forcing the subject to periodically interact with the delay condition every 5-seconds in order to continue their progress would help reduce the contact subjects make with competing activities during delay periods. This condition was later modified in Experiment 3 whereby subjects had to press a button that randomly appeared during the delay to proceed in the lesson. The rationale behind this change was to force subjects to pay closer attention to the screen during the contingent interactive delay and potentially limit their contact with competing activities during the pause. In other words, subjects may still have been able to engage in competing activities during the delay but they also needed to carefully watch the screen the entire time to avoid missing their opportunity to escape the pause and move onto the next problem. One important limitation that was not addressed is the possibility that during the delay subjects could have potentially clicked in the location where the "Click here to go to the next problem" button would eventually appear. While it appears that some subjects took advantage of this flaw (e.g., Subjects 60 & 69) this limitation could easily be addressed by randomly moving the button around the screen. Since the contingent interactive delay conditions were designed to keep subjects from contacting competing activities during delay periods (e.g., checking email, watching T.V., etc.) and at least one of the variations appeared to accomplish this task (see Table 1 & Figure 84) these findings provide evidence that limiting contact with competing activities during an enforced delay isn't necessary for reducing racing behavior with contingent postfeedback delays. Instead, simply withholding progress in a lesson following an incorrect answer is enough to prevent racing on the part of some students.

Perhaps the main limitation of the current investigation is the contingent delays used in these experiments did not discriminate between the different causes for low performance. For example, students may perform poorly on online assessments because instruction is sub-standard or questions are poorly worded. Students encountering a delay due to variables outside of their control are especially likely to express feelings of frustration. This circumstance could ultimately result in instructors delivering more carefully planned instruction to avoid student complaints or instructors dropping contingent delays from their online lessons completely. While this issue was controlled for in the current study by using highly controlled math problems, implementing a contingent delay with other forms of instructional content should be done only when instruction has been demonstrated to be clear and effective. Students could also perform poorly on online assessments because they lack the pre-requisite skills needed to accurately answer the questions presented to them. This could result in students racing more frequently on assessment questions if they believe they are going to end up contacting the delay conditions anyway. However, this set of circumstances could also result in more students spending time learning pre-requisite skills to avoid the contingent delays. The best way to address this issue would be arrange a contingency such that contingent delays only follow fast inaccurate responses on the part of the student. Unfortunately, setting up this type of contingency using the E-learning software currently available on the market would require extensive modifications.

Future investigations into contingent postfeedback delays could focus on the relationship between the amount of response effort required to answer a problem and the length of the delay needed to prevent racing. Although the current study showed both 30-seconds and 60-seconds were long enough to prevent racing for some subjects, the 60-second delay appears to have been significantly more effective given that the intervention was ineffective for only 1 subject out of 40. Other future research in this area might examine the effectiveness of using additional forms of aversive stimuli to setup avoidance contingencies to increase academic performance in an online setting. For example, instead of using contingent delays following incorrect answers, an arrangement could be created where additional questions are contingent on incorrect answers. This type of preparation would result in more practice for students who are struggling with the material and would function as a disincentive for those students who understand the material but are racing through the content. Additionally, future work in this area could examine using contingent postfeedback delays in combination with other proven techniques. For example, contingent postfeedback delays may be more effective when users are exposed to extra instructional content during the delays. Finally, future research should investigate making postfeedback delays contingent on a more specific criteria than incorrect answers. For example, making postfeedback delays contingent on very fast response times and incorrect answers would be a better criterion and would help differentiate between performances where the subject tried to answer correctly but failed and subjects who are simply racing through the lesson as quickly as possible.

REFERENCES

Allen, I. E. & Seaman, J. (2010). *Learning on Demand: Online Education in the United States, 2009.* Babson Survey Research Group.

Allen, I. E. & Seaman, J. (2011). *Going the Distance: Online Education in the United States, 2011.* Babson Survey Research Group.

Bilodeau, E. A. & Bilodeau, I. M. (1958). Variation of temporal intervals among critical events in five studies of knowledge results. *Journal of Experimental Psychology, 55(6)*, 603-612.

Bodemer, D., Ploetzner, R., Feuerlein, I. & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. Learning and Instruction, 14, 325–341.

Bourne, L. E. & Bunderson, C. V. (1963). Effects of delay of informative feedback and length of postfeedback interval on concept identification. *Journal of Experimental Psychology, 65(1)*, 1-5.

Brown, J. S. (1949). A proposed program of research on psychological feedback (knowledge of results) in the performance of psychomotor tasks. In Research Planning Conference on Perceptual and Motor Skills. San Antonio, Texas: Human Resource Research Center, Lackland A.F.B.(Conf. Rep. 49-2)., 81-87.

Chapman, B. (2010). *How Long Does it Take to Create Learning?* Chapman Alliance, LLC.

Chase, J. A. & Houmanfar, R. (2009). The differential effects of elaborate feedback and basic feedback on student performance in a modified, Personalized System of Instruction course. *Journal of Behavior Education, 18,* 245-265.

Clariana, R. B. (2003). The effectiveness of constructed-response and multiple-choice study tasks in computer aided learning. Journal of Educational Computing Research, 28, 395–406.

Clariana, R. B. (2004). An interaction of screen colour and lesson task in CAL. British Journal of Educational Technology, 35, 35–43.

Cooper J. O., Heron, T. E. & Heward, W. L. (2007). *Applied Behavior Analysis, 2nd Edition*. Pearson, NJ.

Crosbie, J. & Kelly, G. (1994). Effects of imposed postfeedback delays in programmed instruction. *Journal of Applied Behavior Analysis, 27(3)*, 483-491.

Denny, M. R., Allard, M., Hall, E. & Rokeach, M. (1960). Supplementary report: Delay of knowledge of results, knowledge of task, and intertrial interval. *Journal of Experimental Psychology, 60(5)*, 327.

Ferster, C. B. & Appel, J. B. (1961). Punishment of S^{Δ} responding in matching to sample by time out from positive reinforcement. *Journal of the Experimental Analysis of Behavior, 4(1)*, 45-56.

Fish, W. W. & Wickersham, L. E. (2009). Best practices for online instructors: Reminders. *The Quarterly Review of Distance Education, 10(3)*, 279-284.

Gibbons, A. S., Robertson, D. J., Duffin, J. R. & Thompson, B. (2001). Effects of administering feedback following extended problem solving. Journal of Educational Computing Research, 25, 417–426.

Green, R. S., Eppler, M. A., Ironsmith, M. & Wuensch, K. L. (2007). Review question formats and web design usability in computer-assisted instruction. British Journal of Educational Technology, 38, 679–686.

Hewett, B. L. & Powers, C. E. (2007). Guest Editors' Introduction: Online teaching and learning: Preparation, development, and organizational communication. *Technical Communication Quarterly 16(1)*, 1-11.

Holland, J. G. & Skinner, B. F. (1961). *The Analysis of Behavior.* New York: McGraw Hill.

Johnson, D. A. & Dickinson, A. (in press). Using postfeedback delays to improve retention of computer-based instruction. *Psychological Record.*

Johnston, J. M. & Pennypacker, H. S. (1993). Strategies and Tactics of Behavioral Research, 2nd Ed. New Jersey: Lawrene Erlbaum Associates.

Johnson, D. A. & Rubin, S. (2011). Effectiveness of interactive computer-based instruction: A review of studies published between 1995-2007. *Journal of Organizational Behavior Management, 31(1)*, 55-94.

Katayama, A. D., Shambaugh, R. N. & Doctor, T. (2005). Promoting knowledge transfer with electronic note taking. Teaching of Psychology, 32, 129–131.

Kelly, G. & Crosbie, J. (1997). Immediate and delayed effects of imposed feedback delays in computerized programmed instruction. *Psychological Record, 47(4)*, 687-698.

Means, B., Toyama, T., Murphy, R., Bakia, M. & Jones, K. (2009). Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies. U.S. Department of Education.

Miller, M. L. & Malott, R. W. (1997). The importance of overt responding in programmed instruction even with added incentives for learning. Journal of Behavior Education, 7, 497–503.

Montazemi, A. R. & Wang, F. (1995). On the effectiveness of a neural network for adaptive external pacing. Journal of Artificial Intelligence in Education, 6, 379–404.

Moreno, R. & Valdez, A. (2005). Cognitive load and learning effects of having students organize pictures and words in multimedia environments: The role of student interactivity and feedback. Educational Technology Research and Development, 53, 35–45.

Morrison, G. R., Ross, S. M., Gopalakrishnan, M. & Casey, J. (1995). The effects of feedback and incentives on achievement in computer-based instruction. Contemporary Educational Psychology, 20, 32–50.

Munson, K. J. & Crosbie, J. (1998). Effects of response cost in computerized programmed instruction. *Psychological Record, 48(2),* 233-250.

Schnackenberg, H. L. & Sullivan, H. J. (2000). Learner control over full and lean computer-based instruction under differing ability levels. Educational Technology Research and Development, 48, 19–35.

Shute, V. J., Gawlick, L. A. & Gluck, K. A. (1998). Effects of practice and learner control on short- and long-term gain and efficiency. Human Factors, 40, 296–310.

Tudor, R. M. (1995). Isolating the effects of active responding in computer-based instruction. Journal of Applied Behavior Analysis, 28, 343–344.

White, R. M. & Schmidt, S. W. (1972). Preresponse intervals versus postinformative feedback intervals in concept identification. *Journal of Experimental Psychology, 94(3)*, 350-352.

Yadrick, R. M., Regian, J. W., Connolly-Gomez, C. & Robertson-Schule, L. (1997). Dyadic vs. individual practice with exploratory and directive mathematics tutors. Journal of Educational Computing Research, 17, 165–186.

Young, A. & Norgard, C. (2006). Assessing the quality of online courses from the students' perspective. *Internet and Higher Education, 9*, 107-115.

Zimmerman, J. & Baydan, N. T. (1963). Punishment of S^{Δ} responding of humans in conditional matching to sample by time-out. *Journal of the Experimental Analysis of Behavior, 6(4)*, 589-597.

Zimmerman, J. & Ferster, C. B. (1963). Intermittent punishment of S^{Δ} responding in matching to sample. *Journal of the Experimental Analysis of Behavior, 6(5)*, 349-356.

Experiment 1: Proficiency Group Subjects

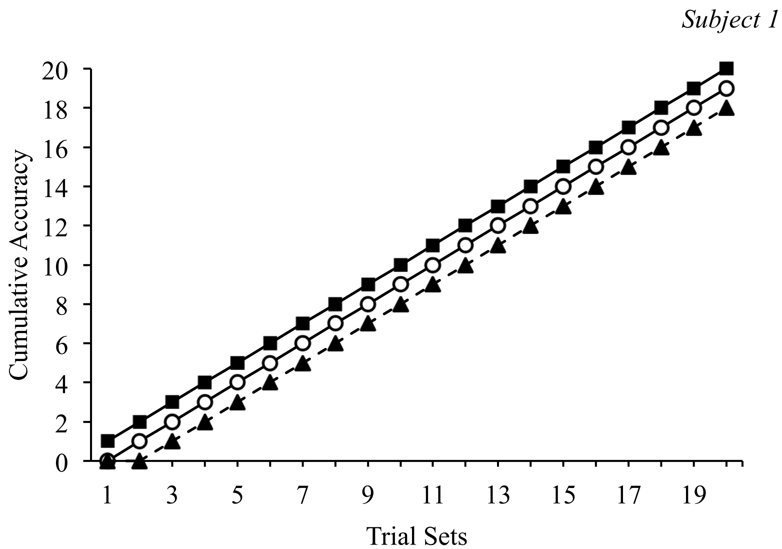


FIGURE 1a. Cumulative accuracy across Control, CD and CID conditions.

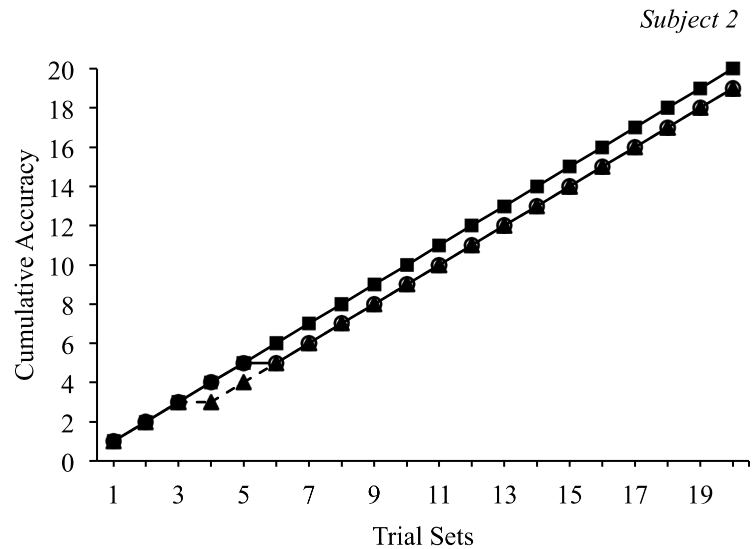


FIGURE 2a. Cumulative accuracy across Control, CD and CID conditions.

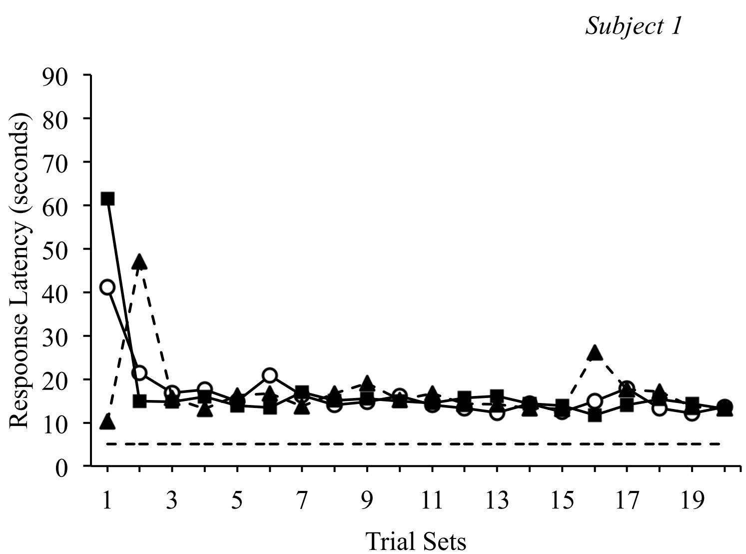


FIGURE 1b. Response latency measures across Control, CD and CID conditions.

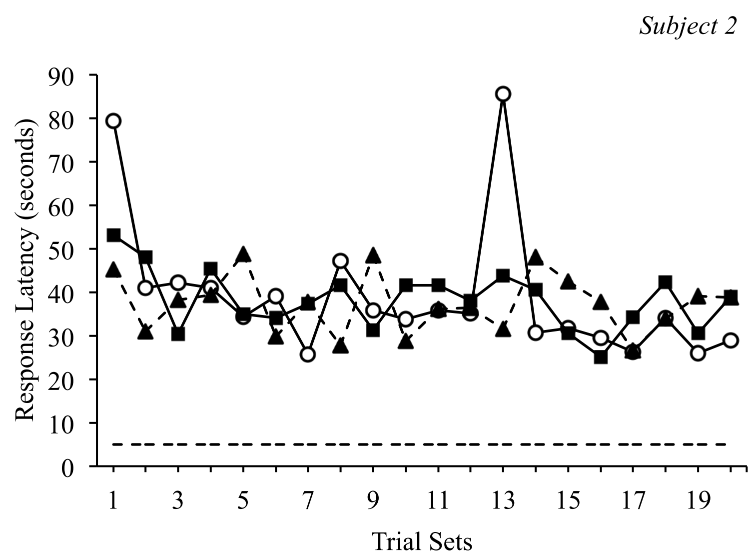


FIGURE 2b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 1: Proficiency Group Subjects

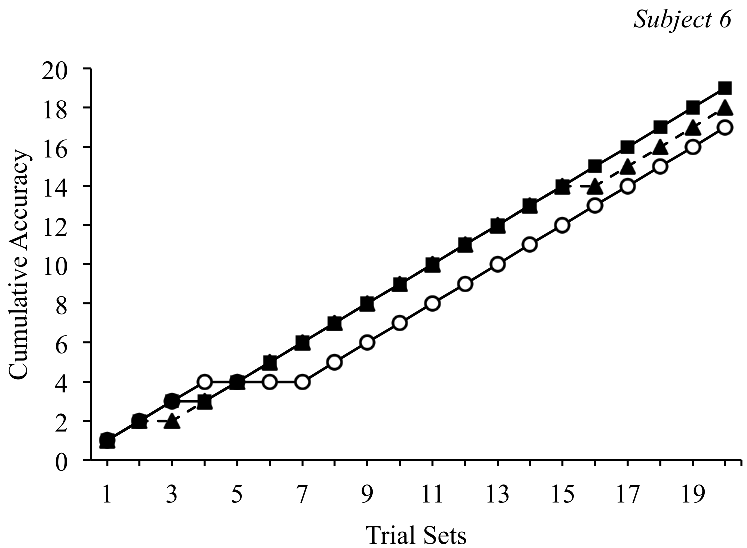


FIGURE 3a. Cumulative accuracy across Control, CD and CID conditions.

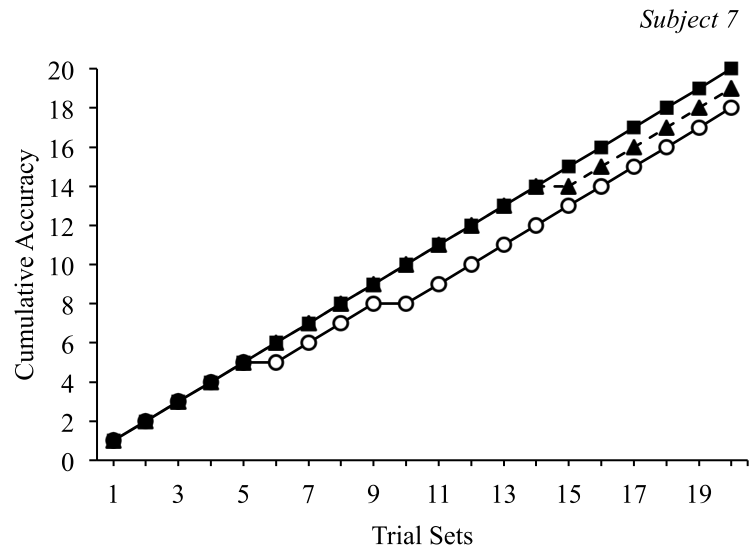


FIGURE 4a. Cumulative accuracy across Control, CD and CID conditions.

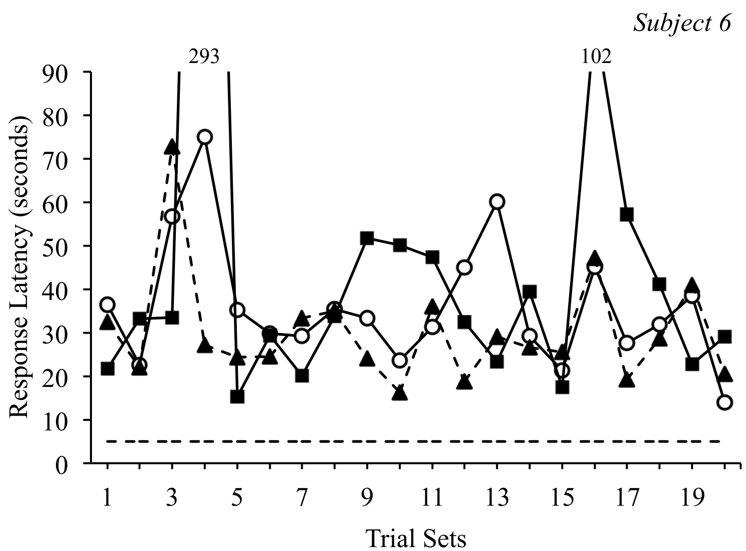


FIGURE 3b. Response latency measures across Control, CD and CID conditions.

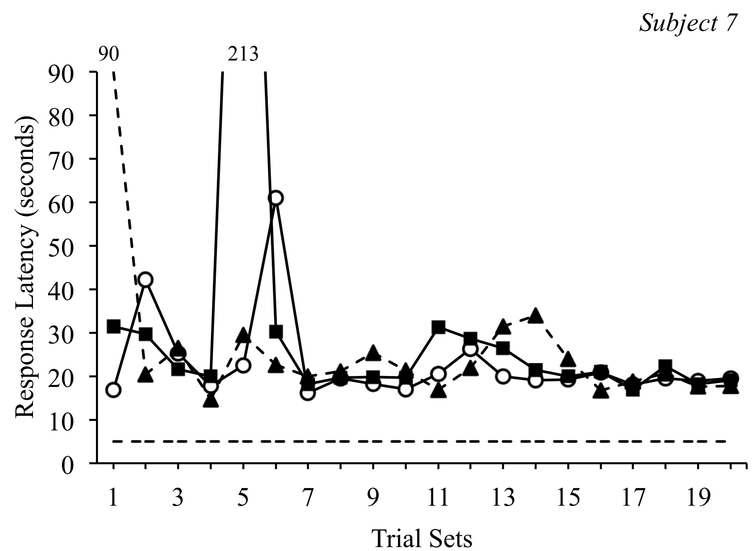


FIGURE 4b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 1: Proficiency Group Subjects

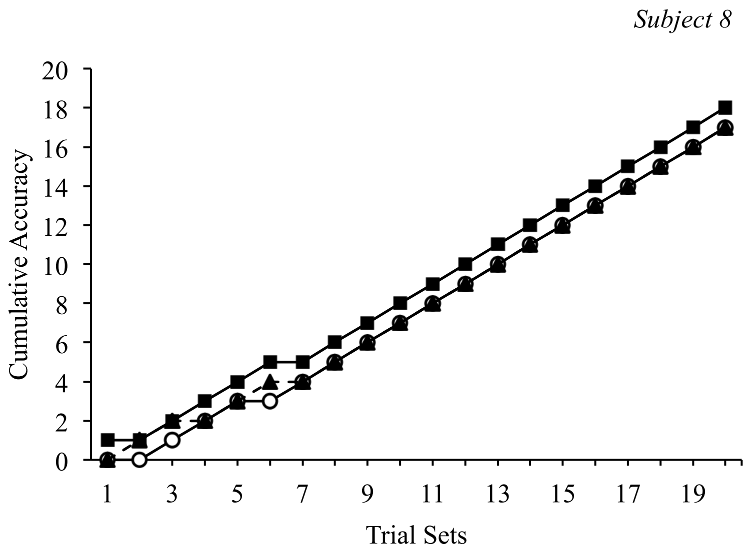


FIGURE 5a. Cumulative accuracy across Control, CD and CID conditions.

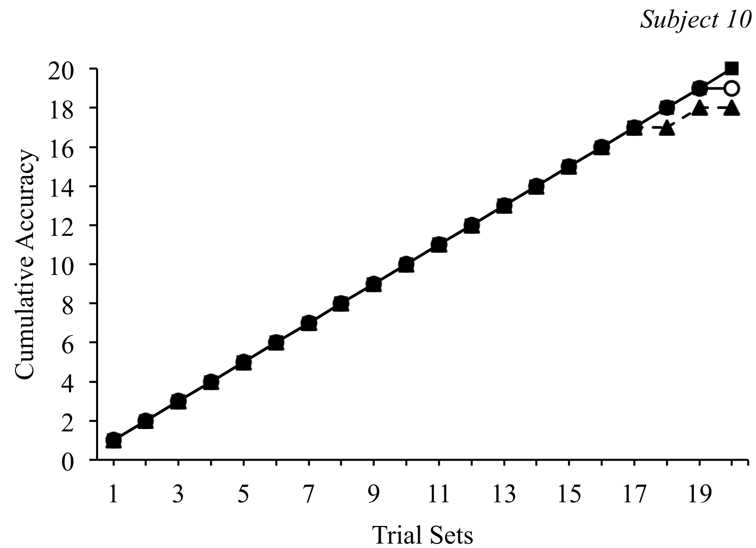


FIGURE 6a. Cumulative accuracy across Control, CD and CID conditions.

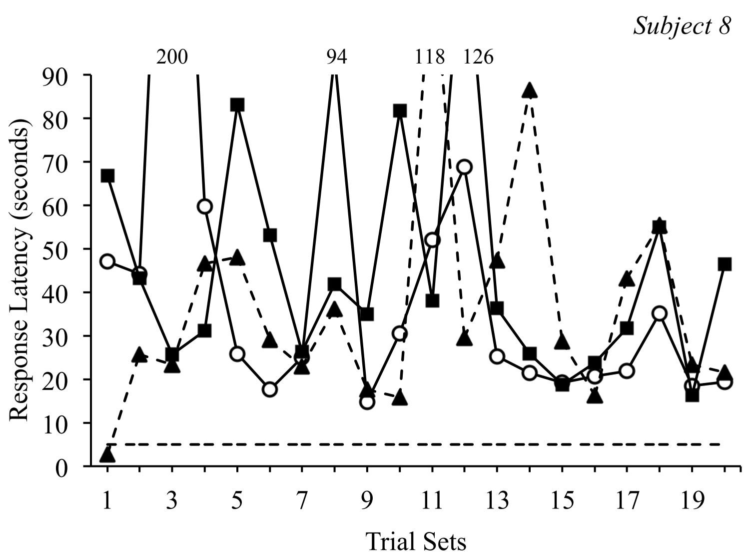


FIGURE 5b. Response latency measures across Control, CD and CID conditions.

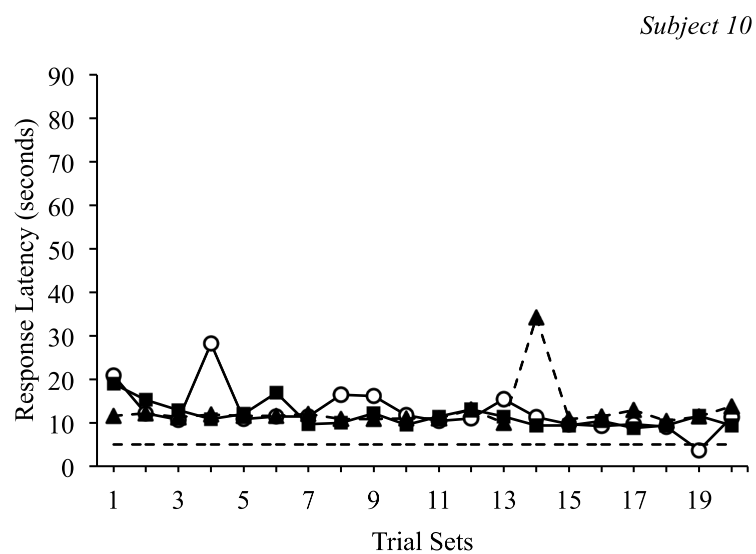


FIGURE 6b. Response latency measures across Control, CD and CID conditions.

Control — Contingent Delay — Contingent Interactive Delay — 5-seconds

Experiment 1: Proficiency Group Subjects

Subject 11

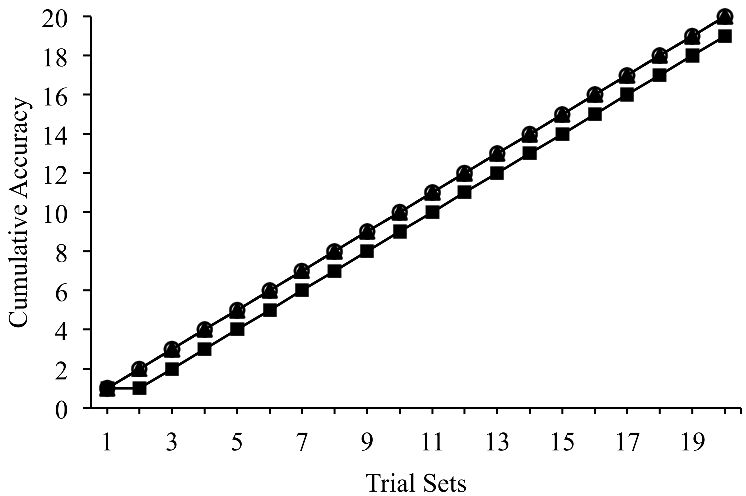


FIGURE 7a. Cumulative accuracy across Control, CD and CID conditions.

Subject 12

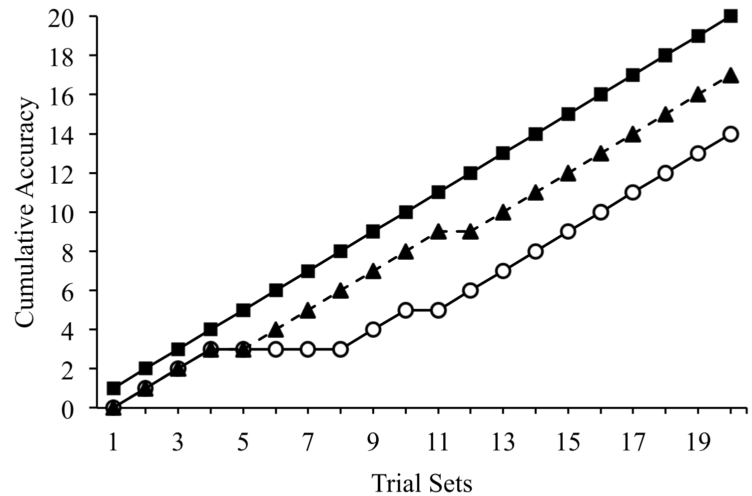


FIGURE 8a. Cumulative accuracy across Control, CD and CID conditions.

Subject 11

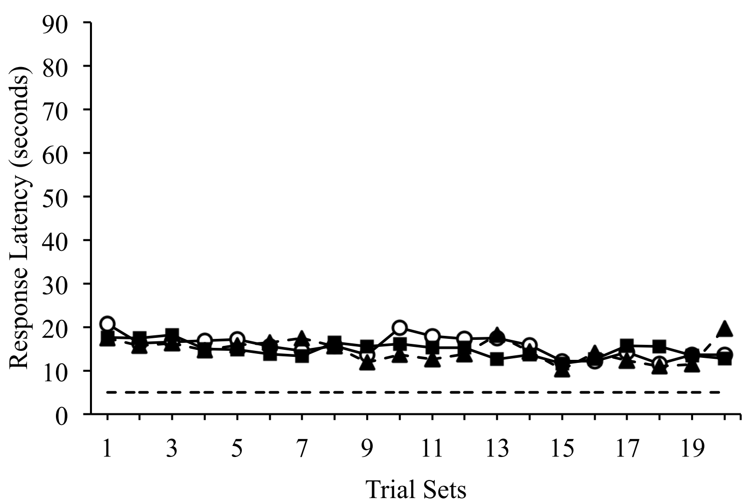


FIGURE 7b. Response latency measures across Control, CD and CID conditions.

Subject 12

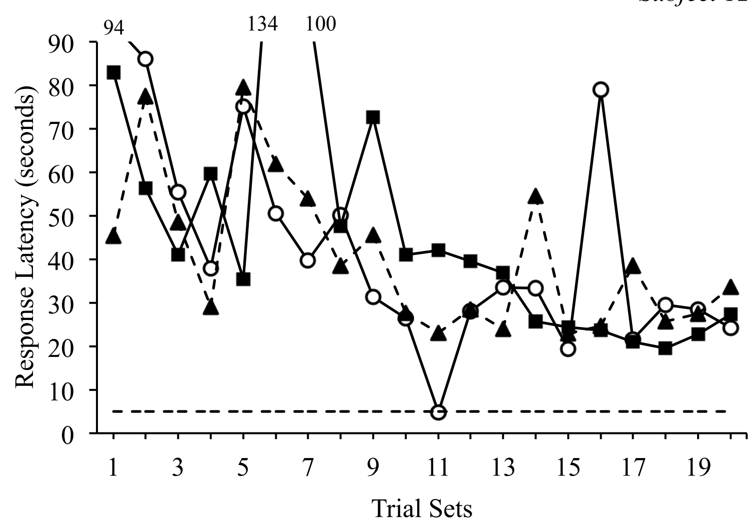


FIGURE 8b. Response latency measures across Control, CD and CID conditions.

Control Contingent Delay Contingent Interactive Delay 5-seconds

Experiment 1: Proficiency Group Subjects

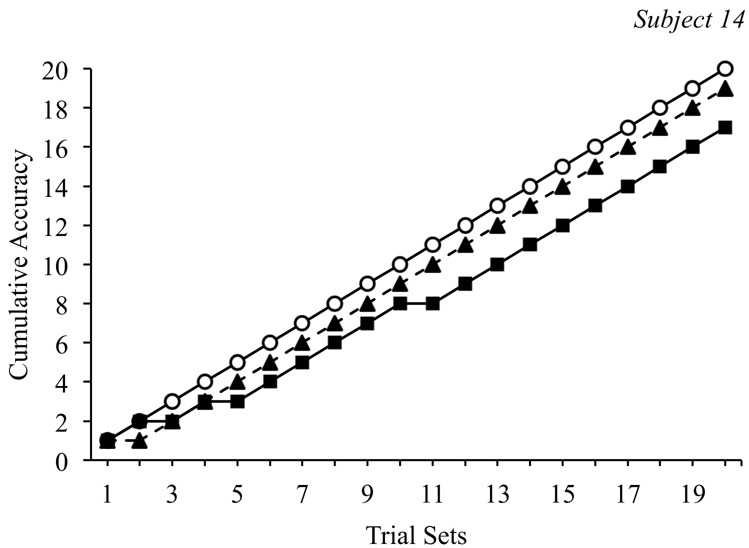


FIGURE 9a. Cumulative accuracy across Control, CD and CID conditions.

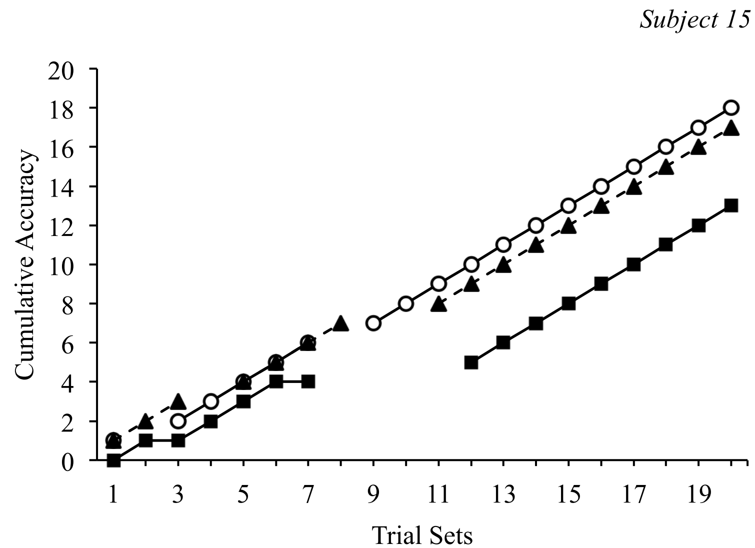


FIGURE 10a. Cumulative accuracy across Control, CD and CID conditions.

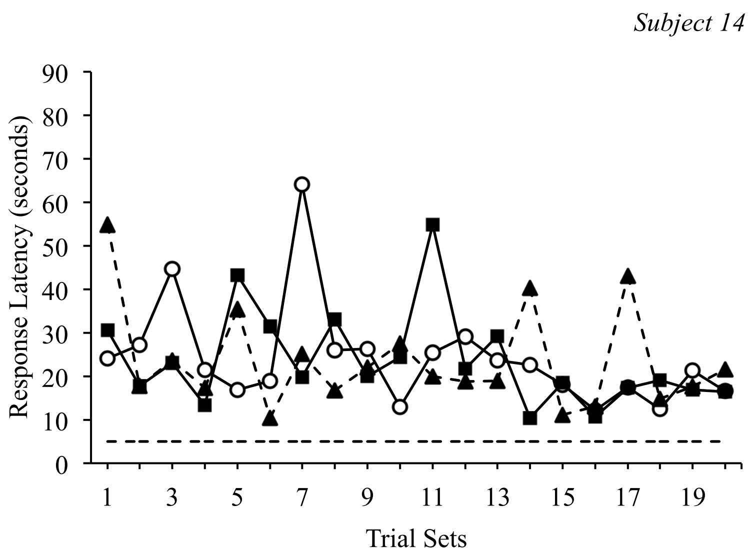


FIGURE 9b. Response latency measures across Control, CD and CID conditions.

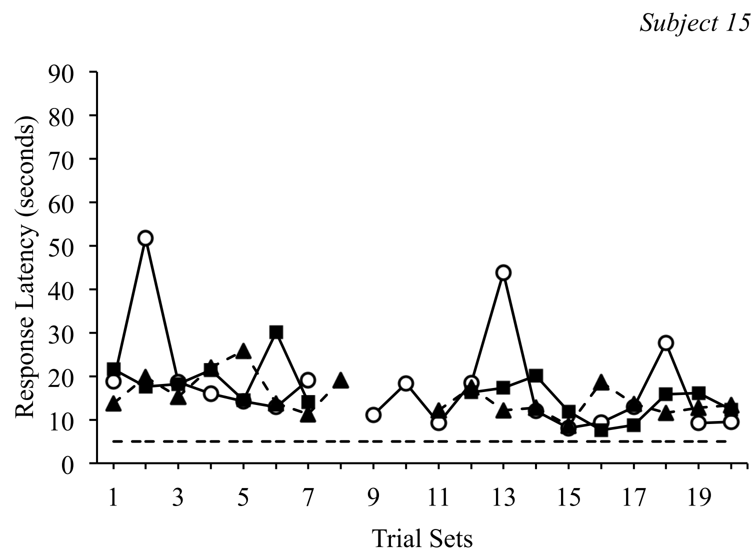


FIGURE 10b. Response latency measures across Control, CD and CID conditions.

Control — Contingent Delay — Contingent Interactive Delay — 5-seconds

Experiment 1: Proficiency Group Subjects

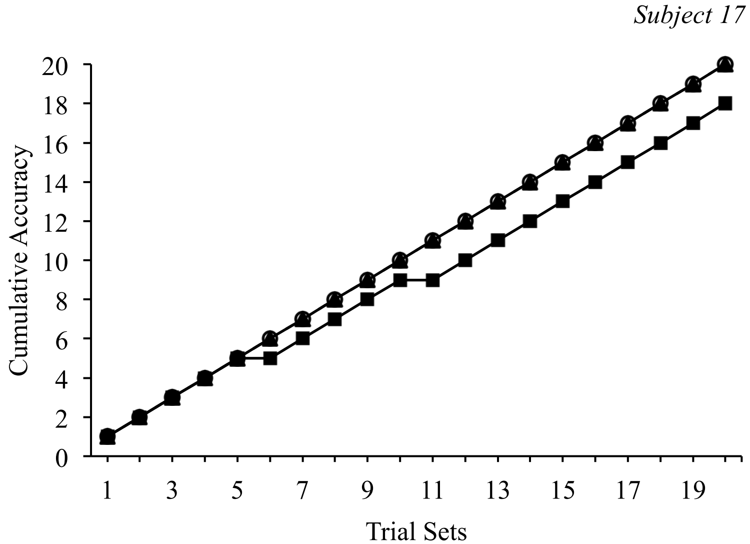


FIGURE 11a. Cumulative accuracy across Control, CD and CID conditions.

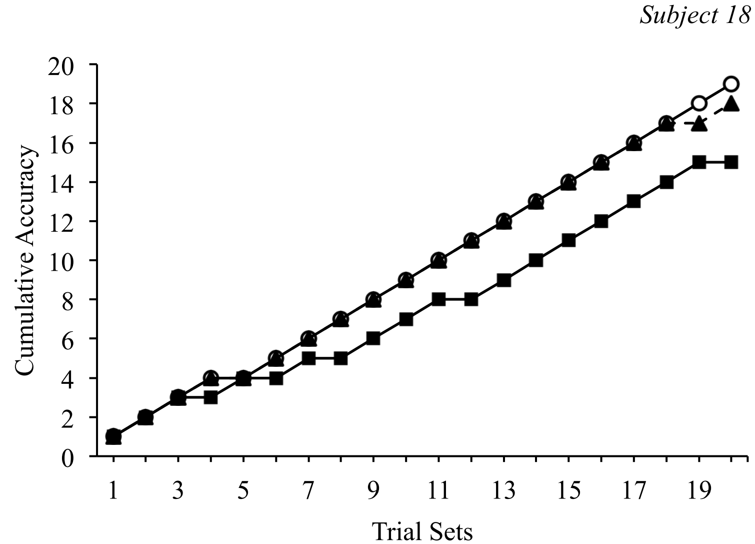


FIGURE 12a. Cumulative accuracy across Control, CD and CID conditions.

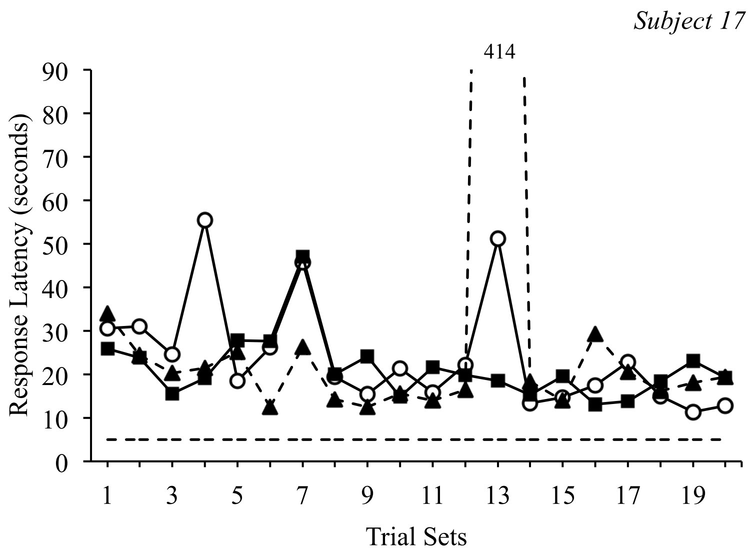


FIGURE 11b. Response latency measures across Control, CD and CID conditions.

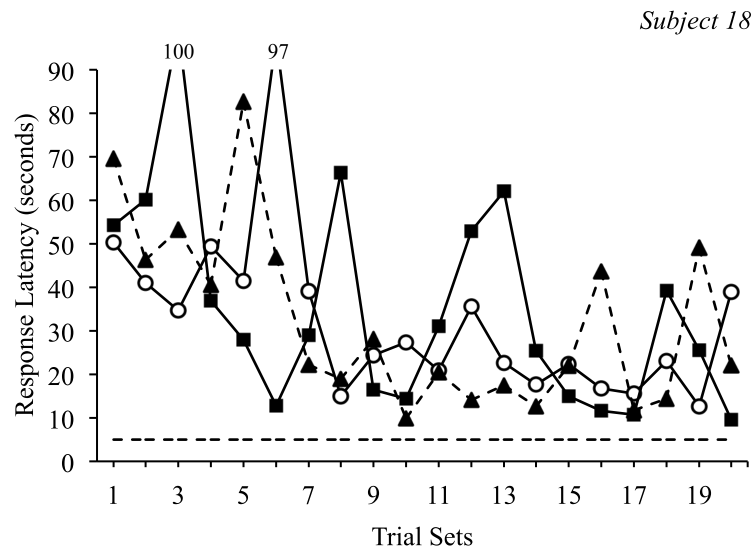


FIGURE 12b. Response latency measures across Control, CD and CID conditions.

Control | Contingent Delay | Contingent Interactive Delay | 5-seconds

Experiment 1: Proficiency Group Subjects

Subject 19

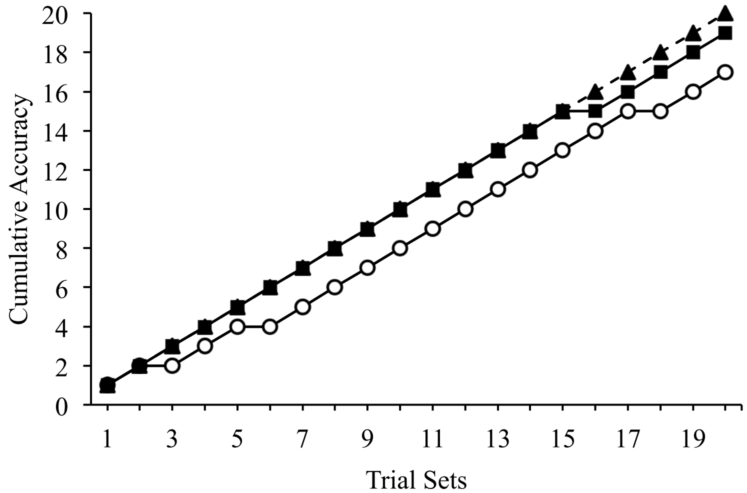


FIGURE 13a. Cumulative accuracy across Control, CD and CID conditions.

Subject 20

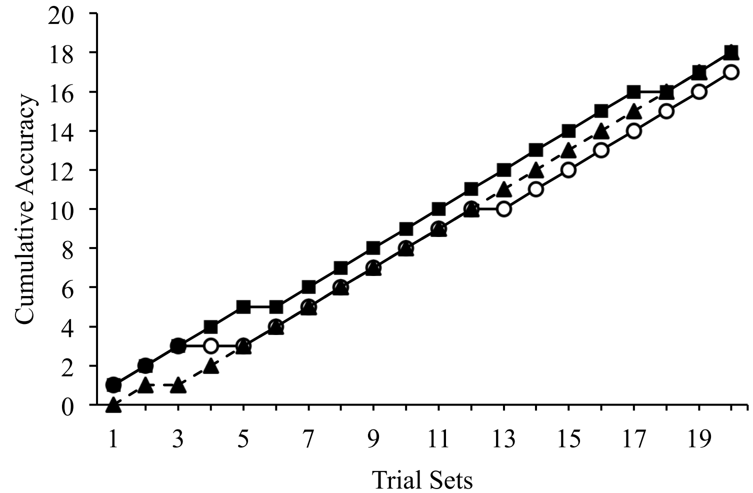


FIGURE 14a. Cumulative accuracy across Control, CD and CID conditions.

Subject 19

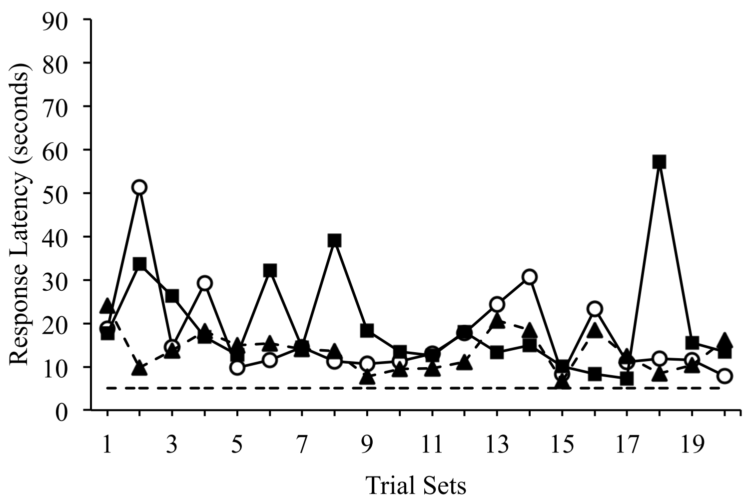


FIGURE 13b. Response latency measures across Control, CD and CID conditions.

Subject 20

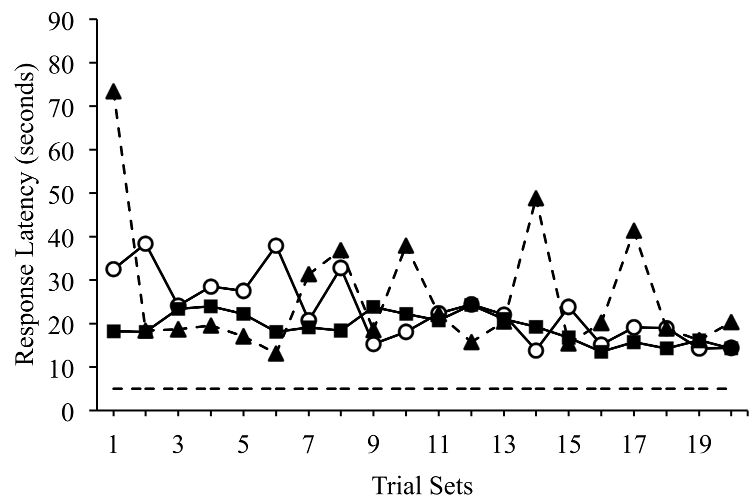


FIGURE 14b. Response latency measures across Control, CD and CID conditions.

Control — Contingent Delay — Contingent Interactive Delay — 5-seconds

Experiment 1: Proficiency Group Subjects

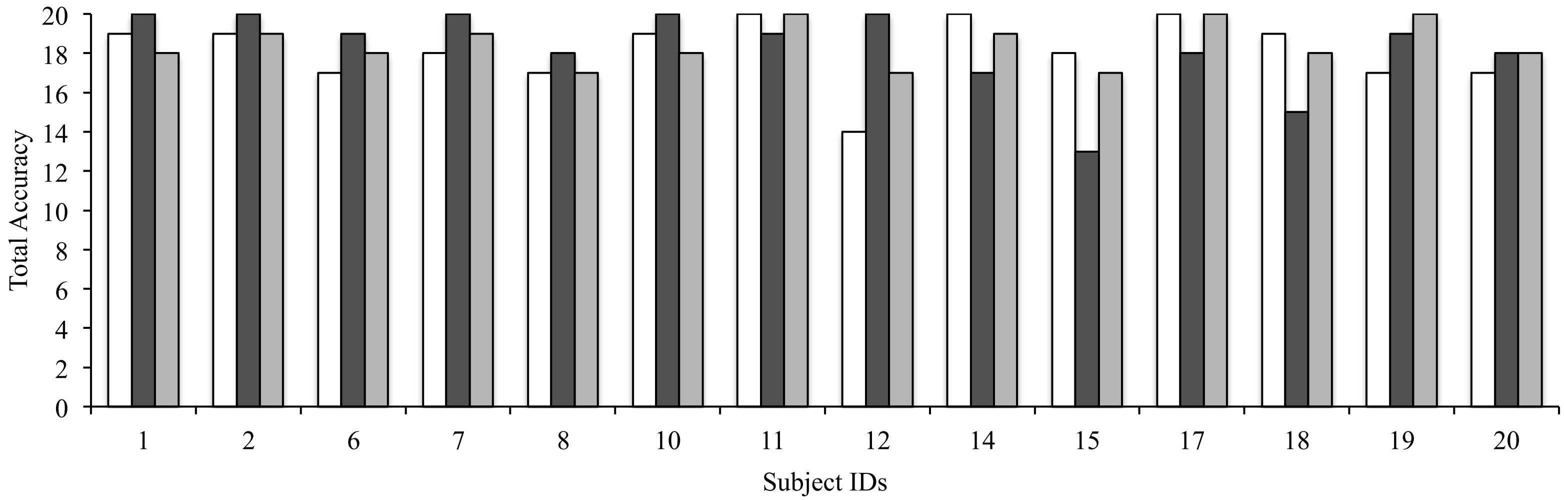


FIGURE 15a. Total accuracy of responding for all Subjects in the Proficiency Group across Control, CD and CID conditions.

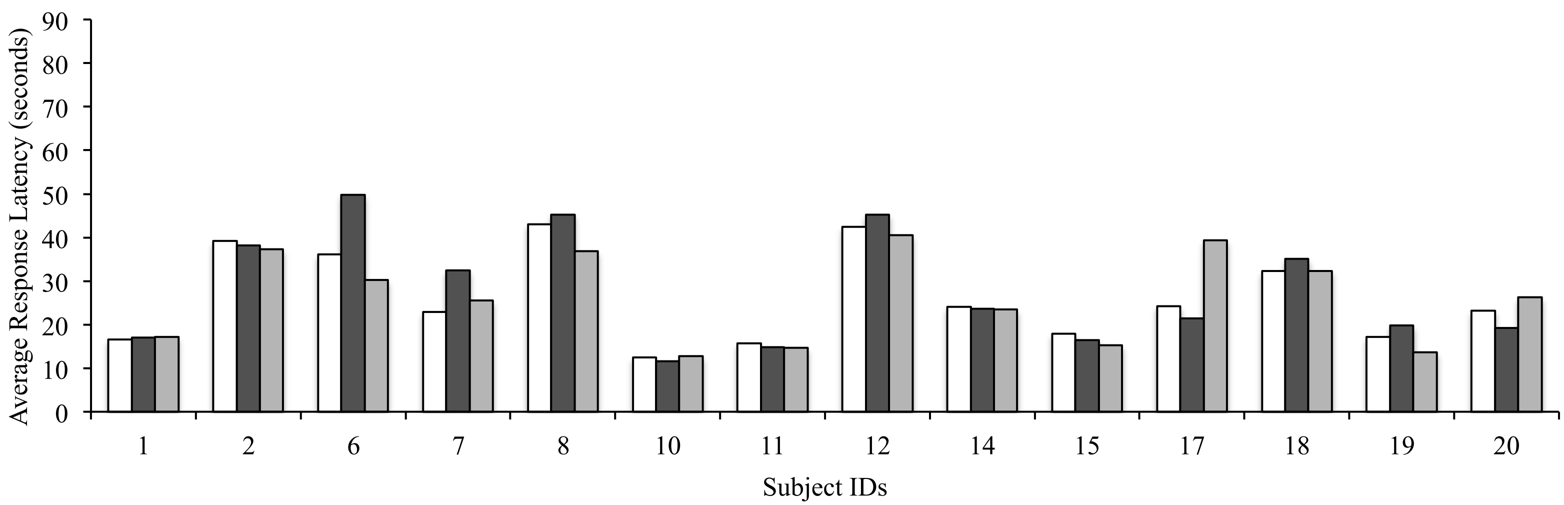


FIGURE 15b. Average response latency measures for all Subjects in the Proficiency Group across Control, CD and CID conditions.

☐ Control ■ Contingent Delay ■ Contingent Interactive Delay

Experiment 1: Avoidance Group Subjects

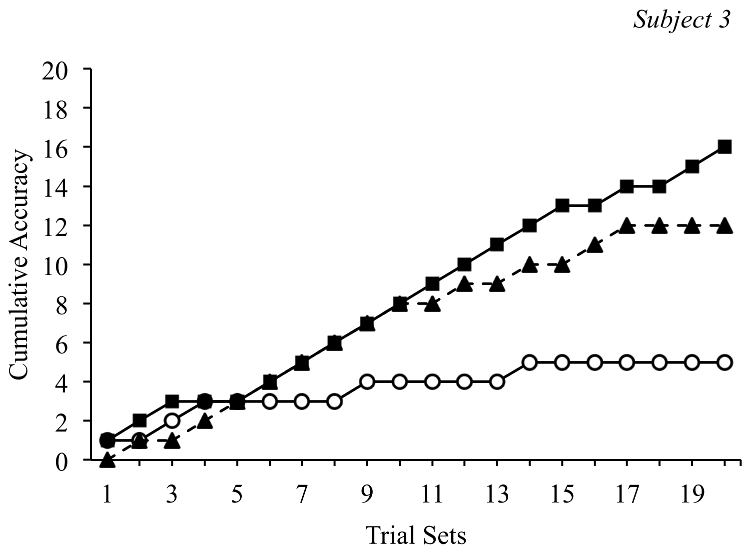


FIGURE 16a. Cumulative accuracy across Control, CD and CID conditions.

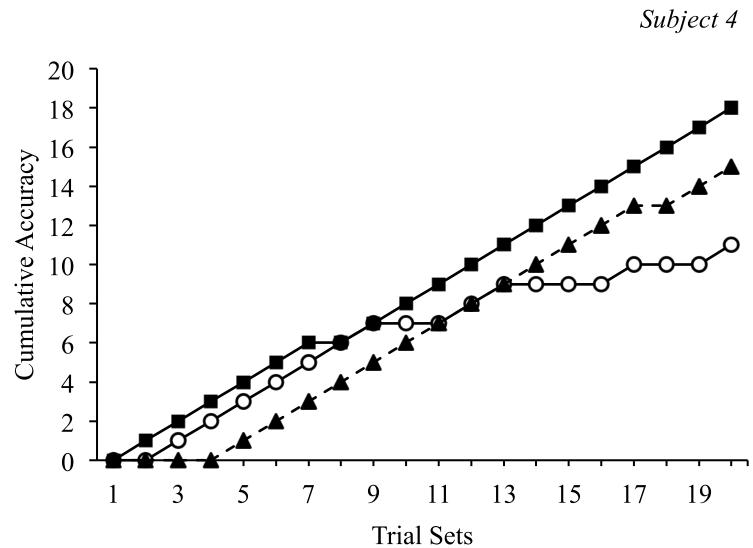


FIGURE 17a. Cumulative accuracy across Control, CD and CID conditions.

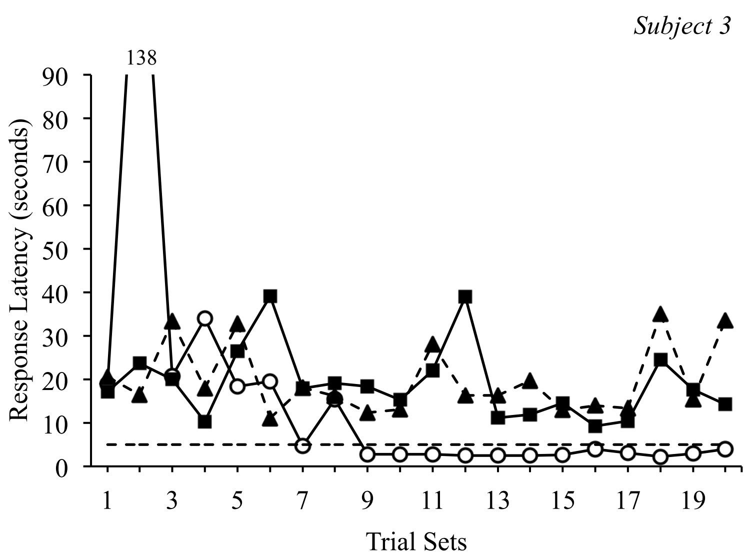


FIGURE 16b. Response latency measures across Control, CD and CID conditions.

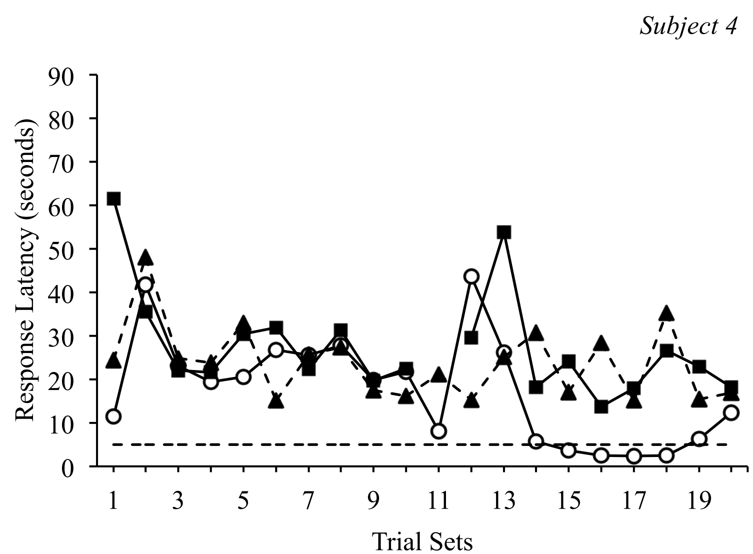


FIGURE 17b. Response latency measures across Control, CD and CID conditions.

Control Contingent Delay Contingent Interactive Delay 5-seconds

Experiment 1: Avoidance Group Subjects

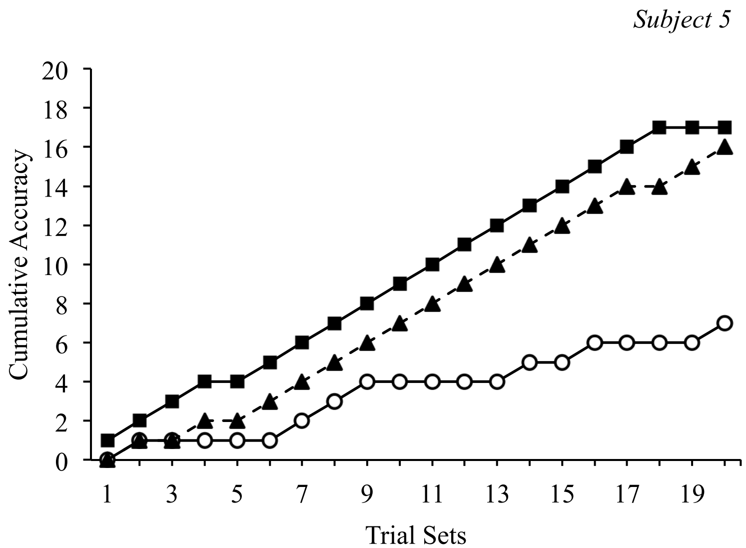


FIGURE 18a. Cumulative accuracy across Control, CD and CID conditions.

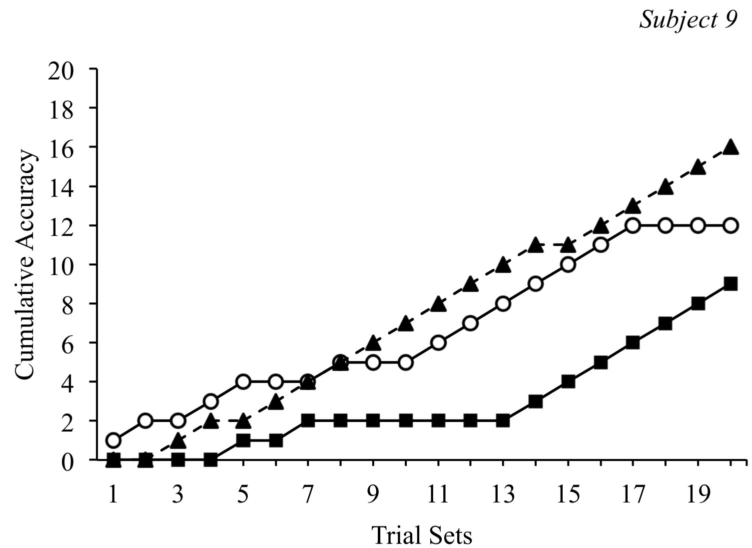


FIGURE 19a. Cumulative accuracy across Control, CD and CID conditions.

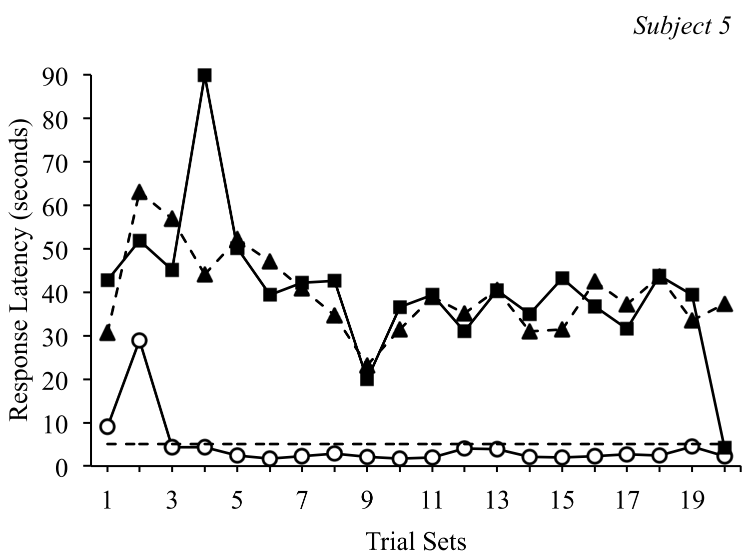


FIGURE 18b. Response latency measures across Control, CD and CID conditions.

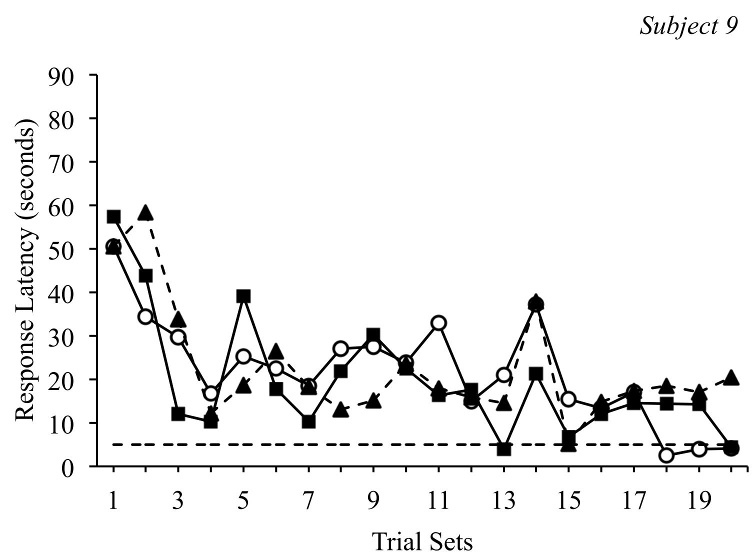


FIGURE 19b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 1: Avoidance Group Subjects

Subject 13

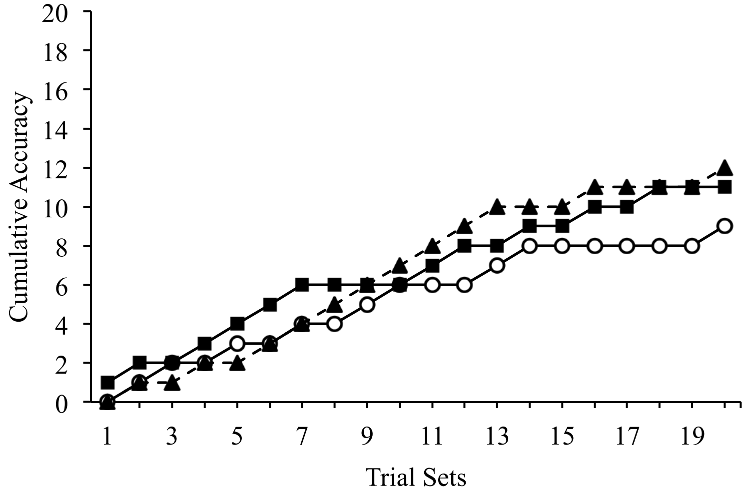


FIGURE 20a. Cumulative accuracy across Control, CD and CID conditions.

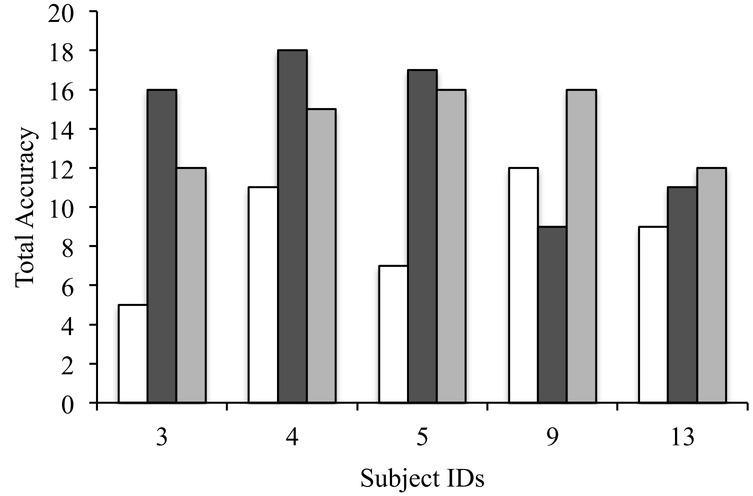


FIGURE 21a. Total accuracy of responding for all Subjects in the Avoidance Group across Control, CD and CID conditions.

Subject 13

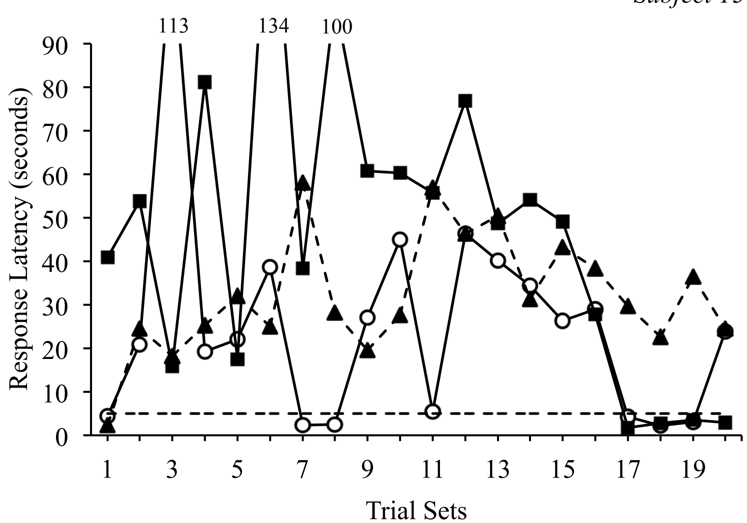


FIGURE 20b. Response latency measures across Control, CD and CID conditions.

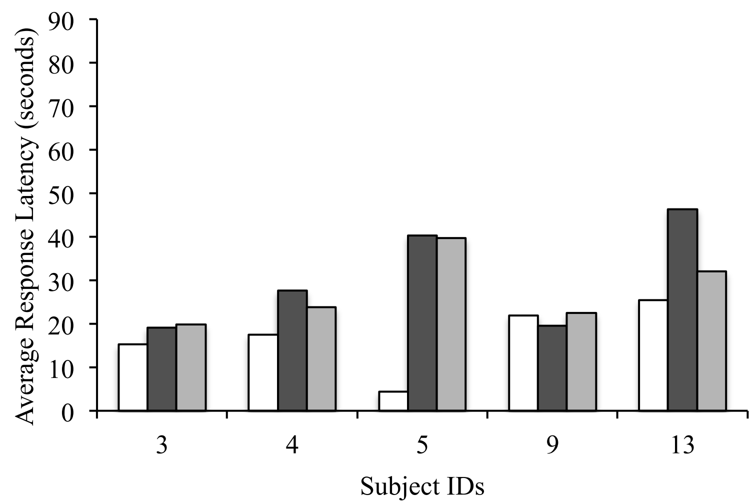


FIGURE 21b. Average response latency measures for all Subjects in the Avoidance Group across Control, CD and CID conditions.

-○- Control -■- Contingent Delay -▲- Contingent Interactive Delay --- 5-seconds

□ Control ■ Contingent Delay ■ Contingent Interactive Delay

Experiment 1: Endurance Group Subject

Subject 16

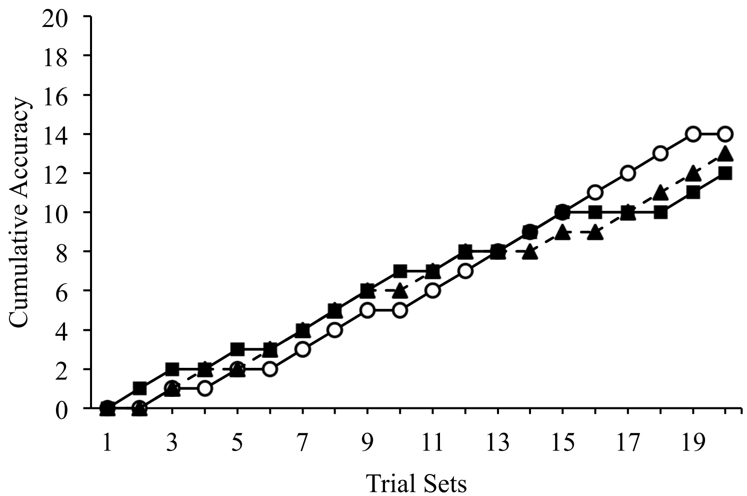


FIGURE 22a. Cumulative accuracy across Control, CD and CID conditions.

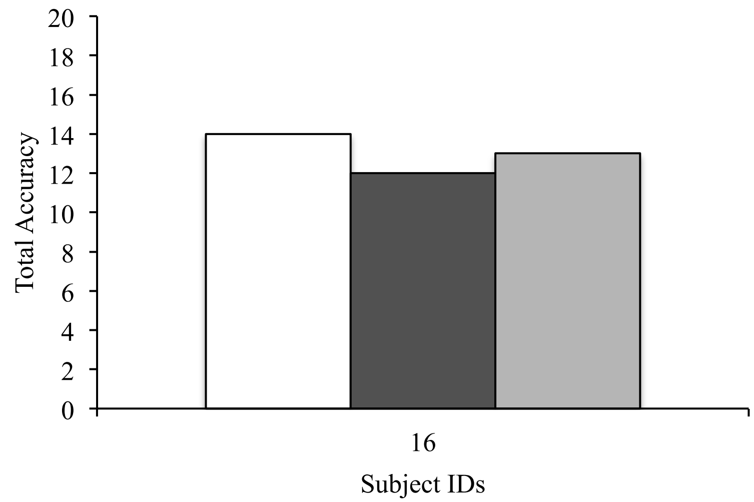


FIGURE 23a. Total accuracy of responding for all Subjects in the Endurance Group across Control, CD and CID conditions.

Subject 16

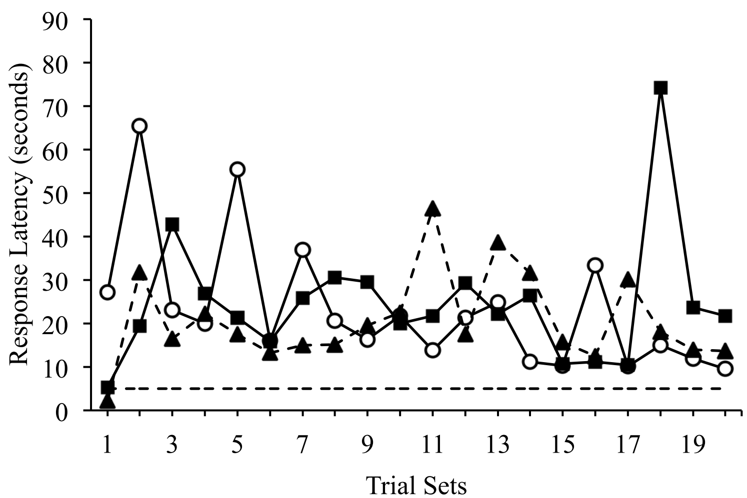


FIGURE 22b. Response latency measures across Control, CD and CID conditions.

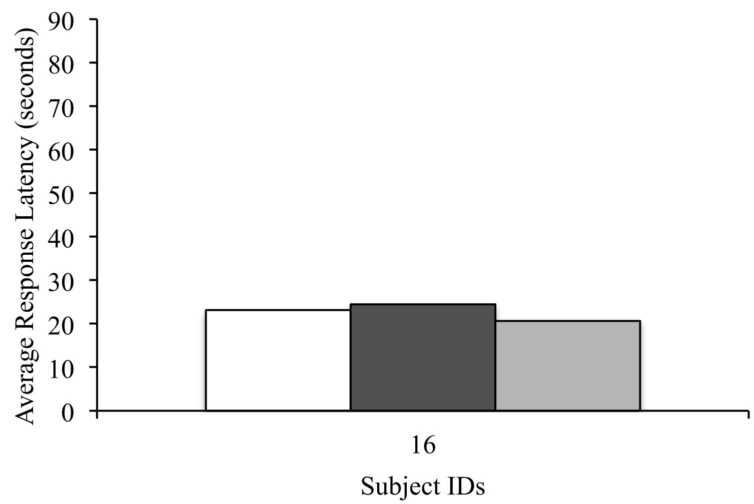


FIGURE 23b. Average response latency measures for all Subjects in the Endurance Group across Control, CD and CID conditions.

Control — Contingent Delay — Contingent Interactive Delay — 5-seconds

Control — Contingent Delay — Contingent Interactive Delay

Experiment 2: Proficiency Group Subjects

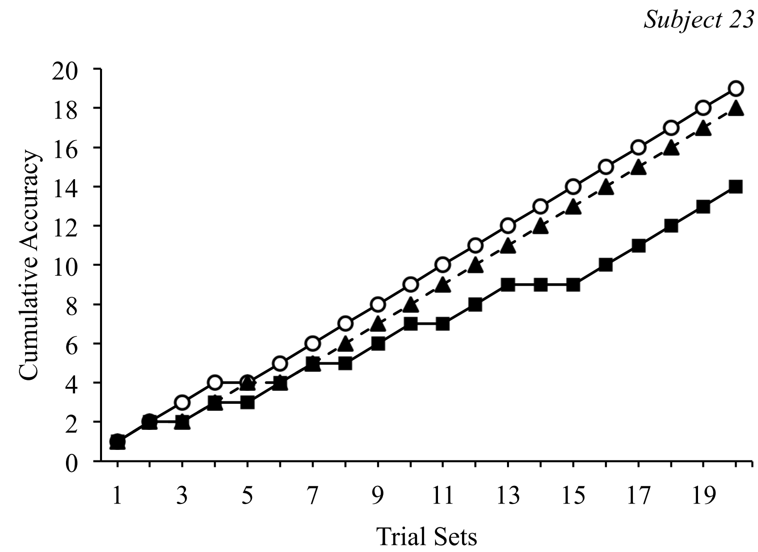


FIGURE 24a. Cumulative accuracy across Control, CD and CID conditions.

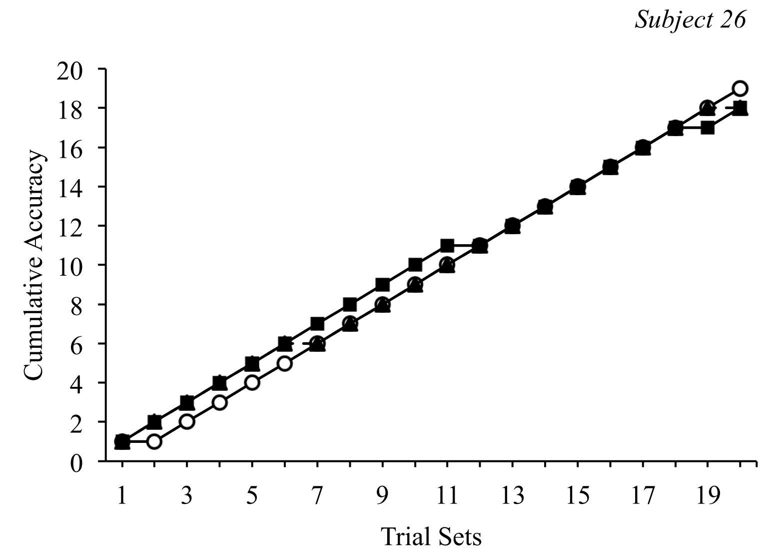


FIGURE 25a. Cumulative accuracy across Control, CD and CID conditions.

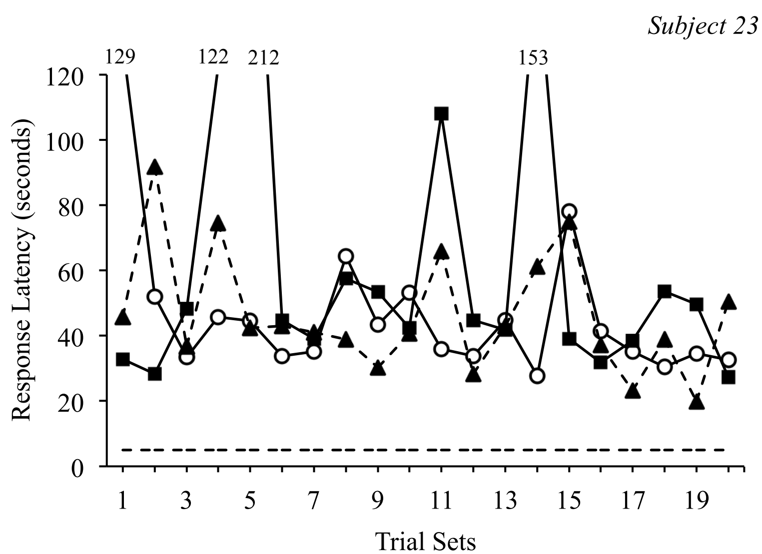


FIGURE 24b. Response latency measures across Control, CD and CID conditions.

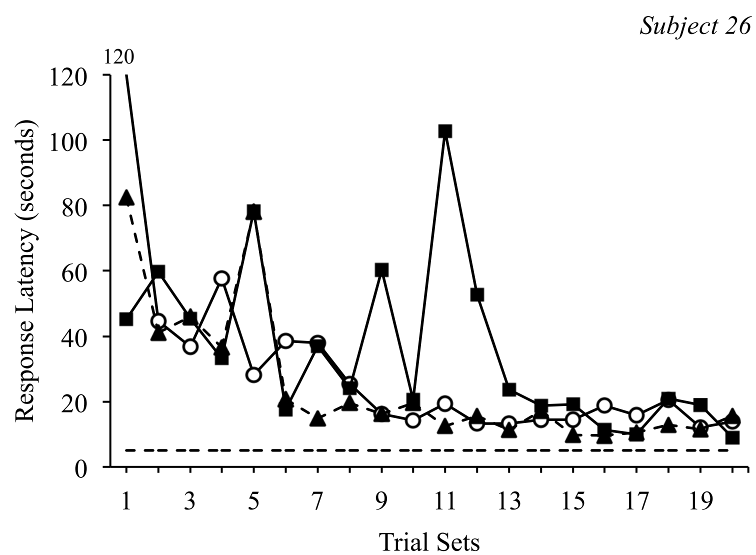


FIGURE 25b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 2: Proficiency Group Subjects

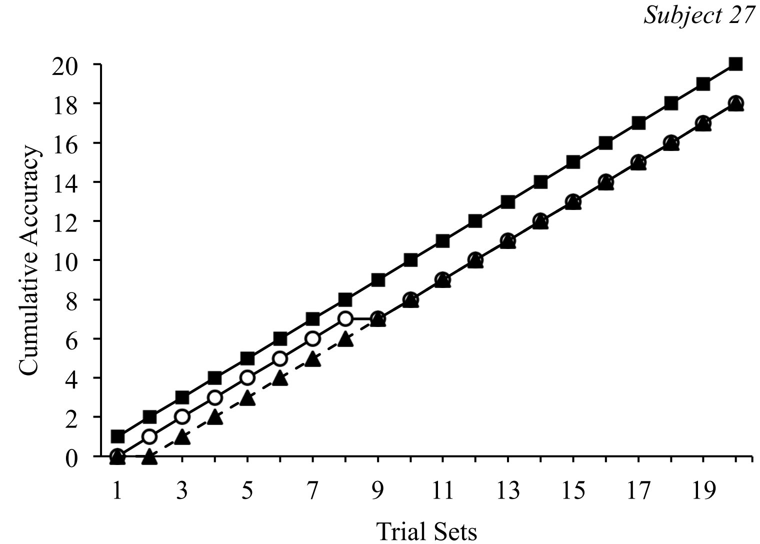


FIGURE 26a. Cumulative accuracy across Control, CD and CID conditions.

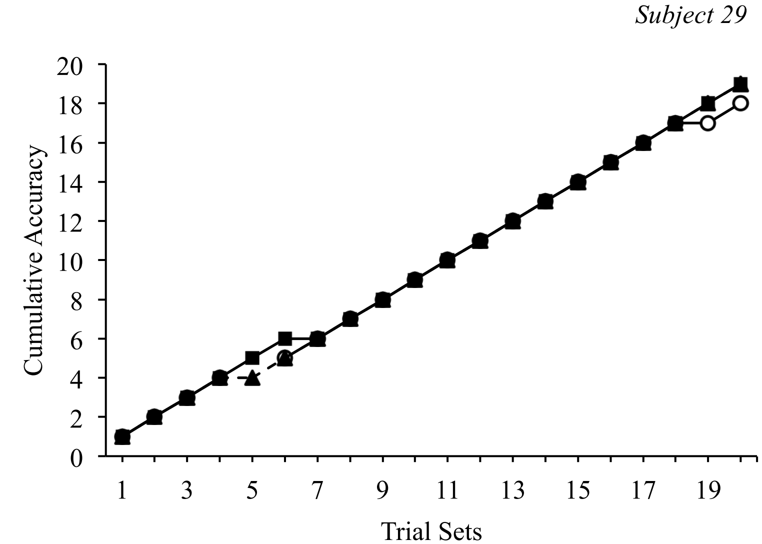


FIGURE 27a. Cumulative accuracy across Control, CD and CID conditions.

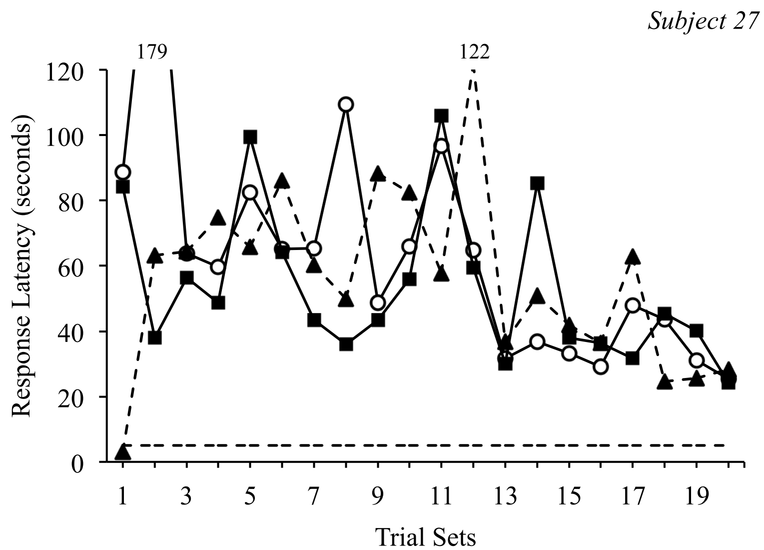


FIGURE 26b. Response latency measures across Control, CD and CID conditions.

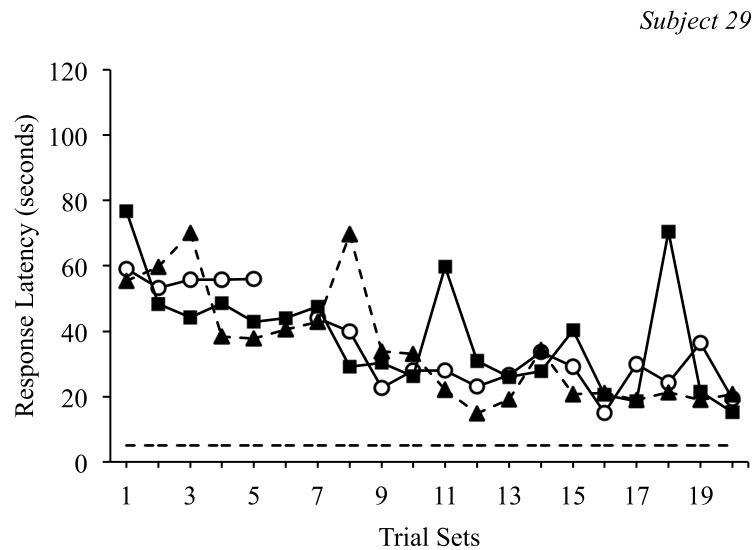


FIGURE 27b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 2: Proficiency Group Subjects

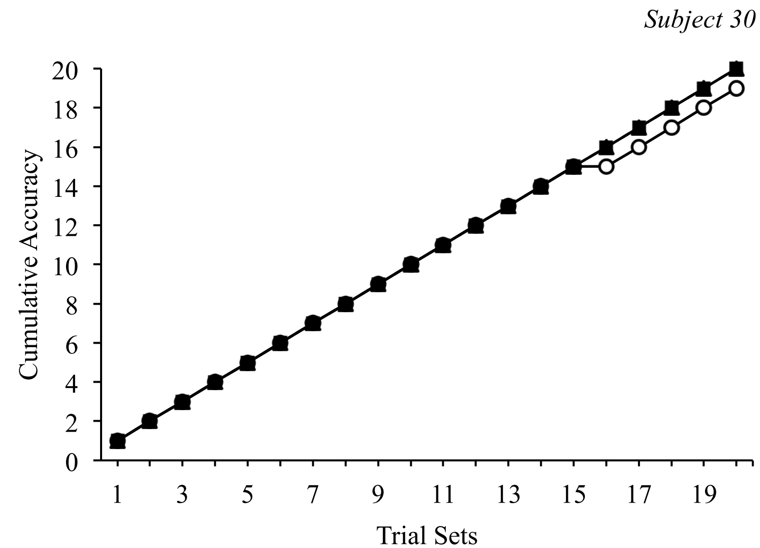


FIGURE 28a. Cumulative accuracy across Control, CD and CID conditions.

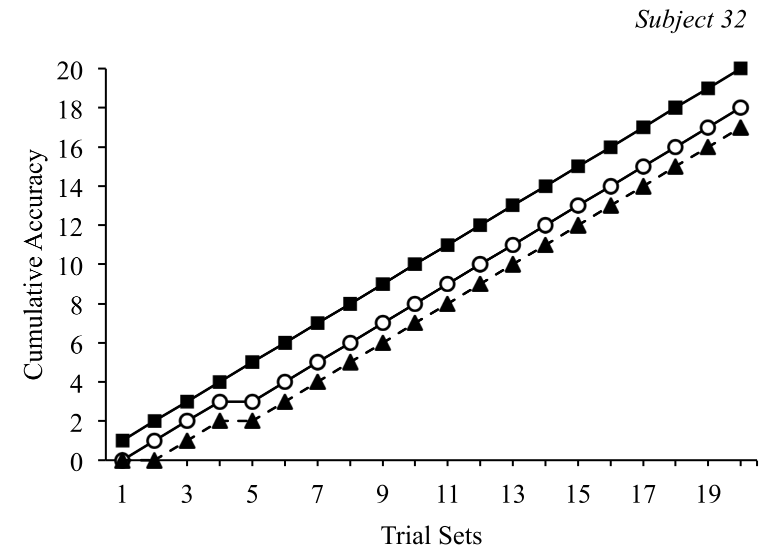


FIGURE 29a. Cumulative accuracy across Control, CD and CID conditions.

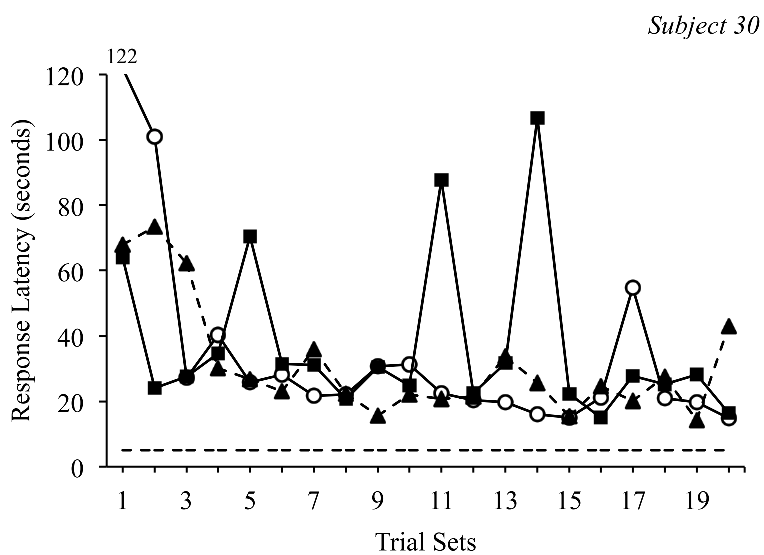


FIGURE 28b. Response latency measures across Control, CD and CID conditions.

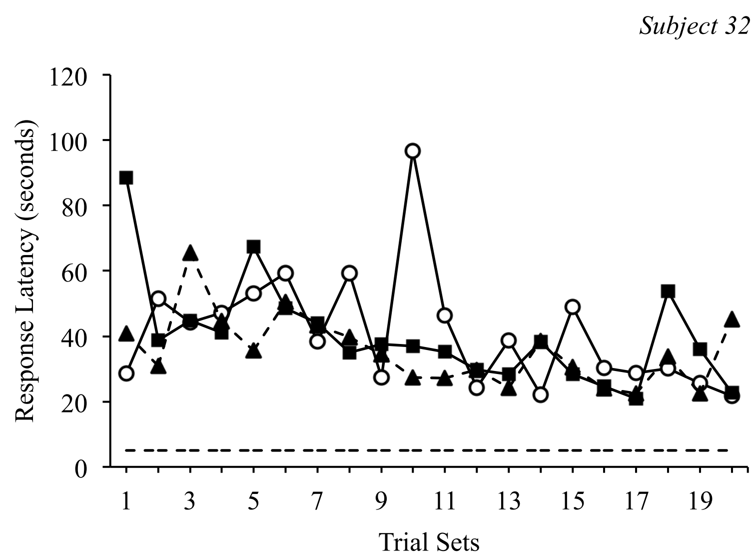


FIGURE 29b. Response latency measures across Control, CD and CID conditions.

Control Contingent Delay Contingent Interactive Delay 5-seconds

Experiment 2: Proficiency Group Subjects

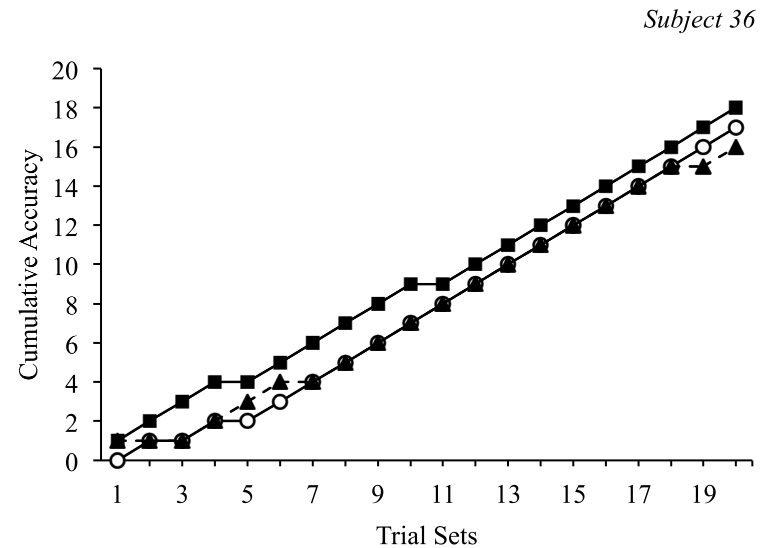


FIGURE 30a. Cumulative accuracy across Control, CD and CID conditions.

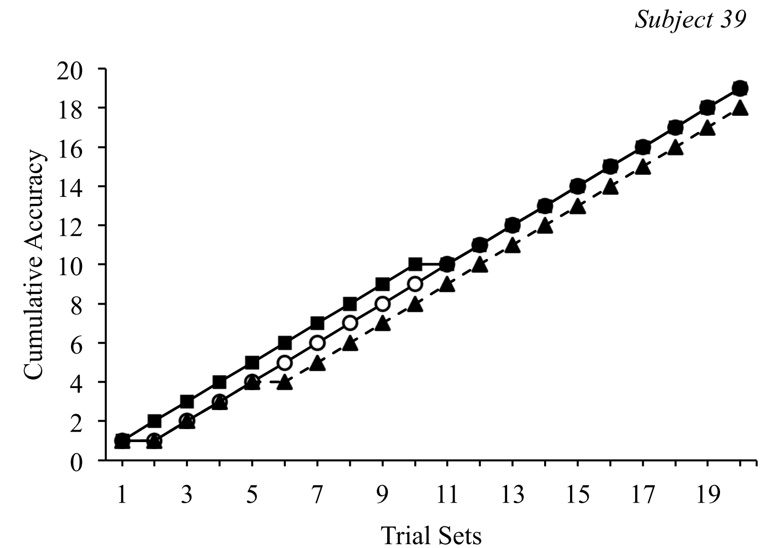


FIGURE 31a. Cumulative accuracy across Control, CD and CID conditions.

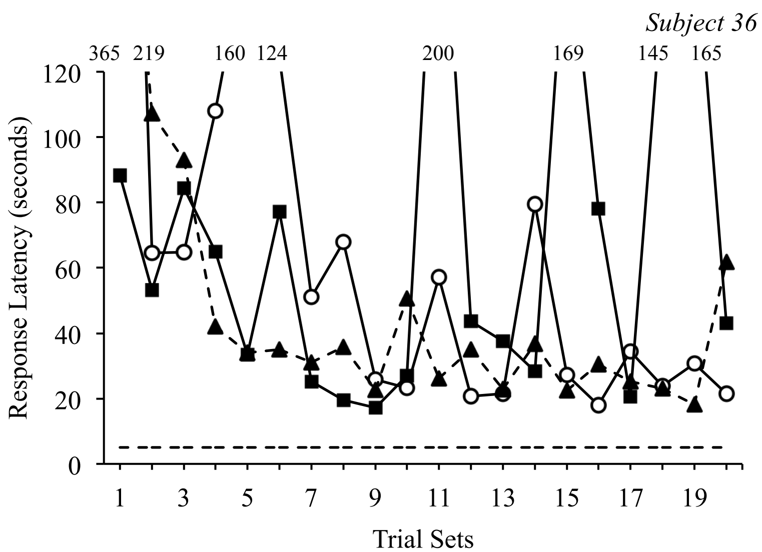


FIGURE 30b. Response latency measures across Control, CD and CID conditions.

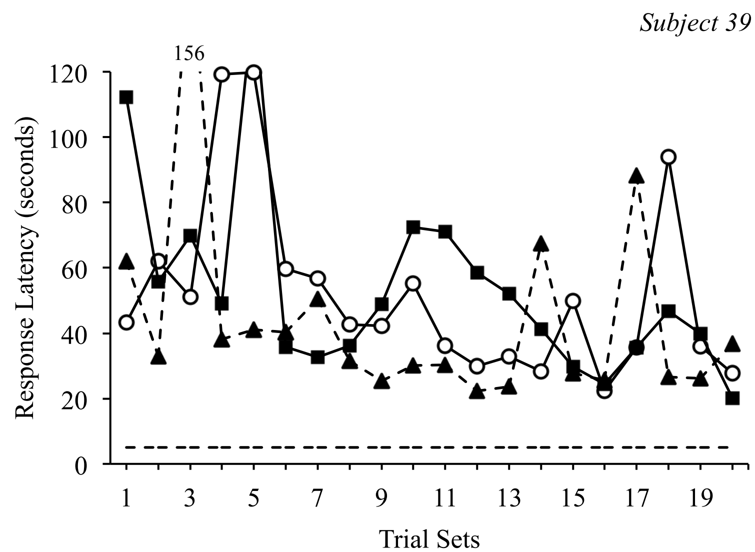


FIGURE 31b. Response latency measures across Control, CD and CID conditions.

Experiment 2: Proficiency Group Subjects

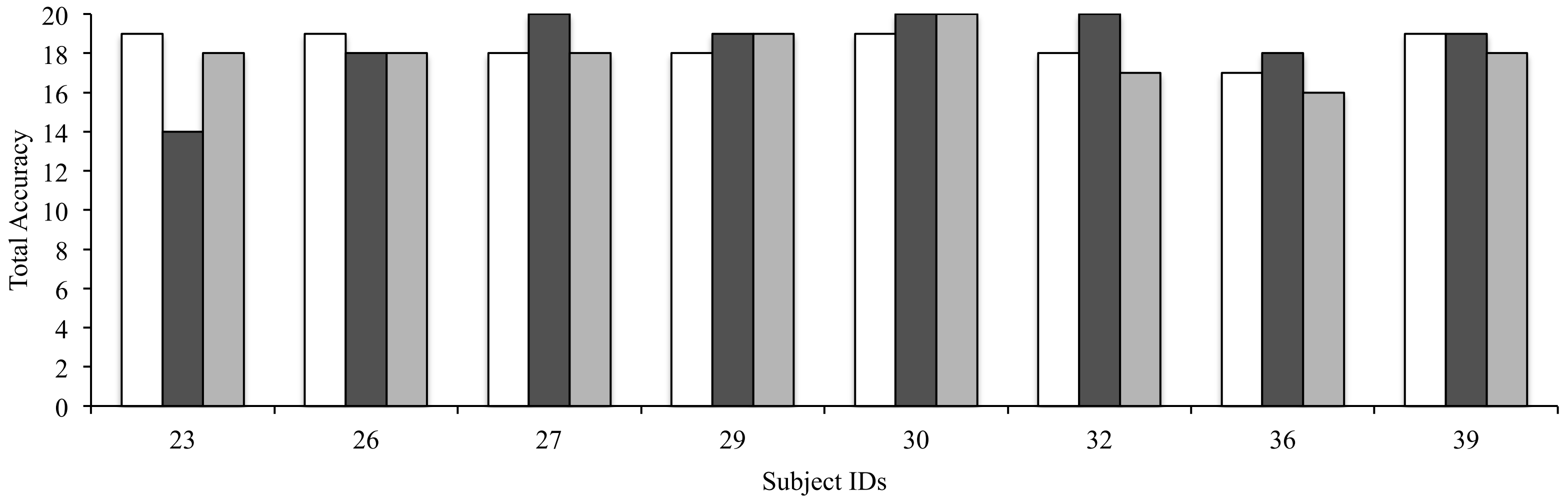


FIGURE 32a. Total accuracy of responding for all Subjects in the Proficiency Group across Control, CD and CID conditions.

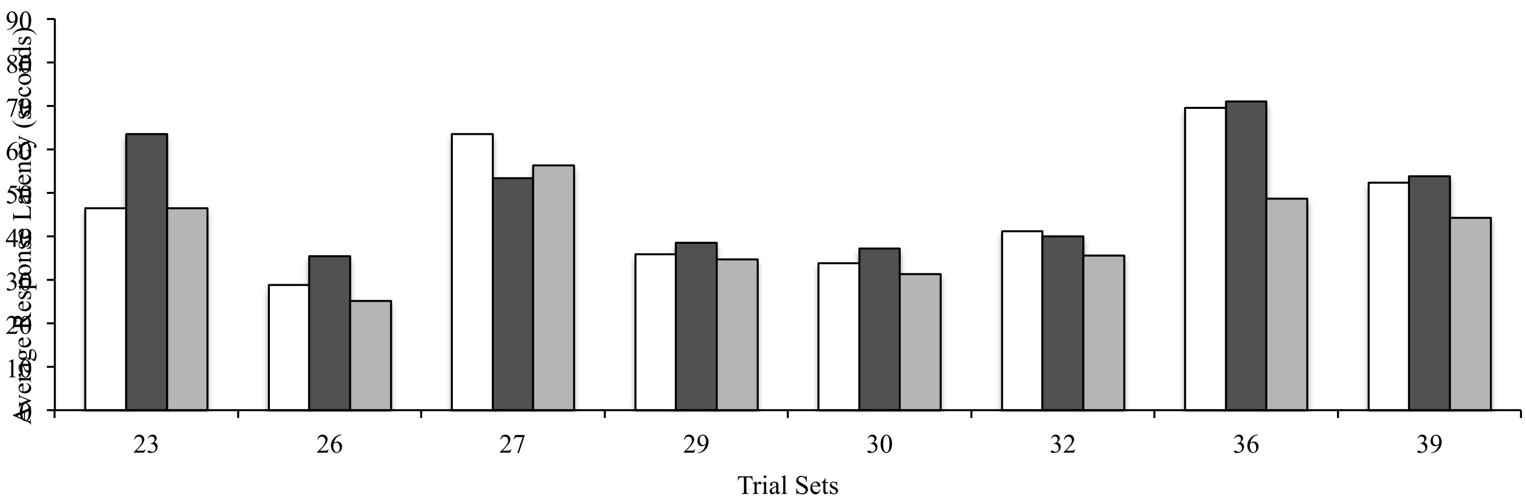


FIGURE 32b. Average response latency measures for all Subjects in the Proficiency Group across Control, CD and CID conditions.

☐ Control ■ Contingent Delay ■ Contingent Interactive Delay

Experiment 2: Avoidance Group Subjects

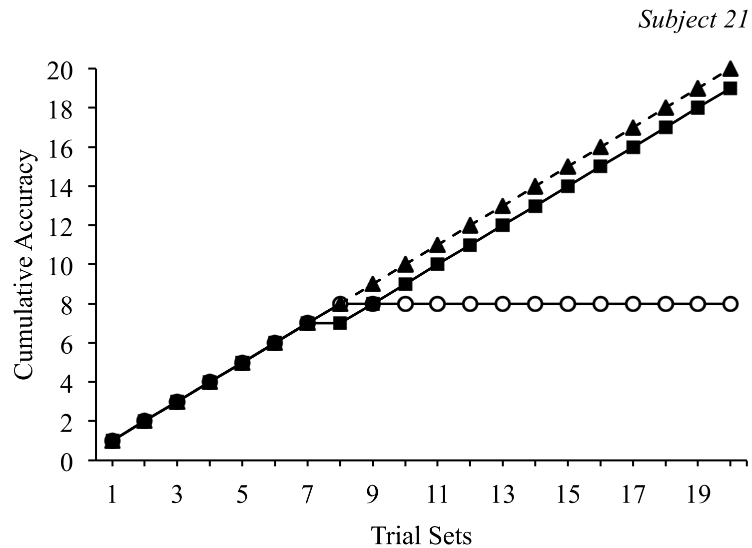


FIGURE 33a. Cumulative accuracy across Control, CD and CID conditions.

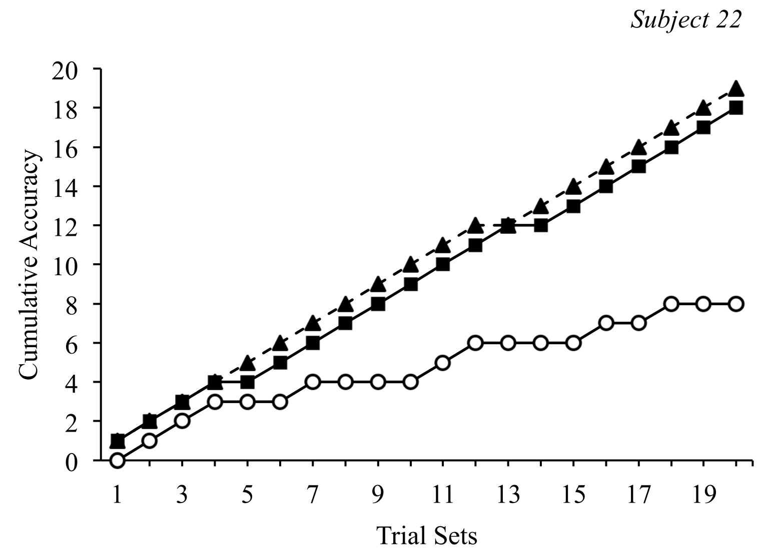


FIGURE 34a. Cumulative accuracy across Control, CD and CID conditions.

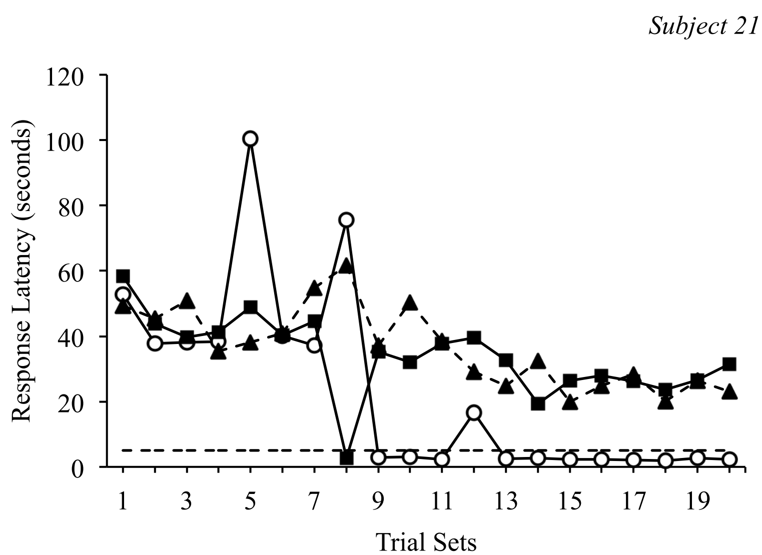


FIGURE 33b. Response latency measures across Control, CD and CID conditions.

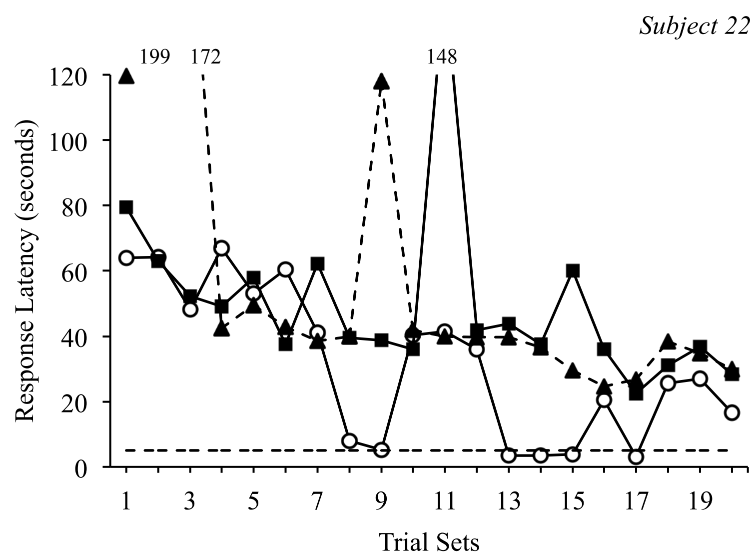


FIGURE 34b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 2: Avoidance Group Subjects

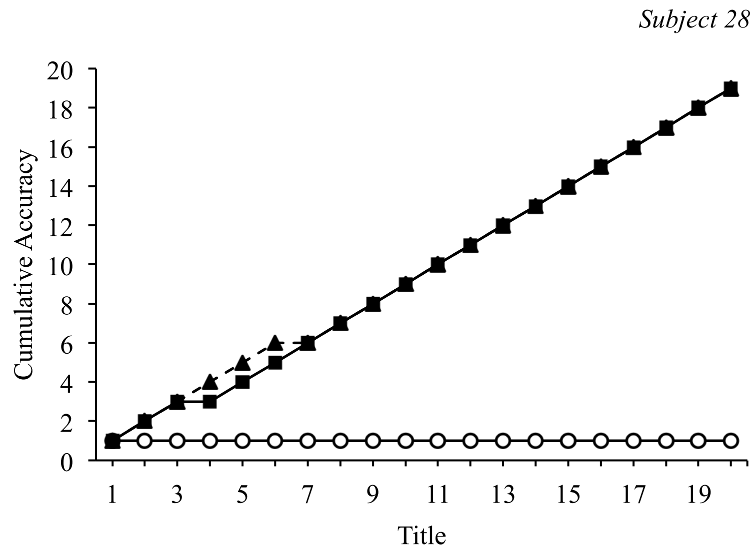


FIGURE 35a. Cumulative accuracy across Control, CD and CID conditions.

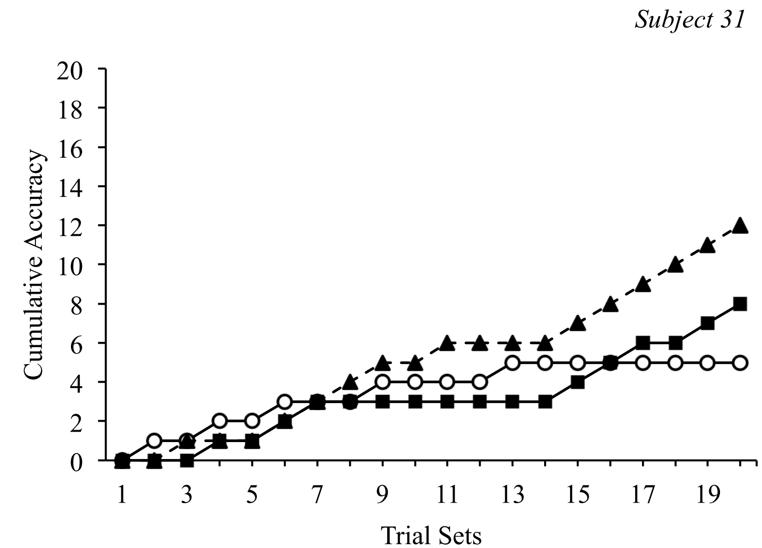


FIGURE 36a. Cumulative accuracy across Control, CD and CID conditions.

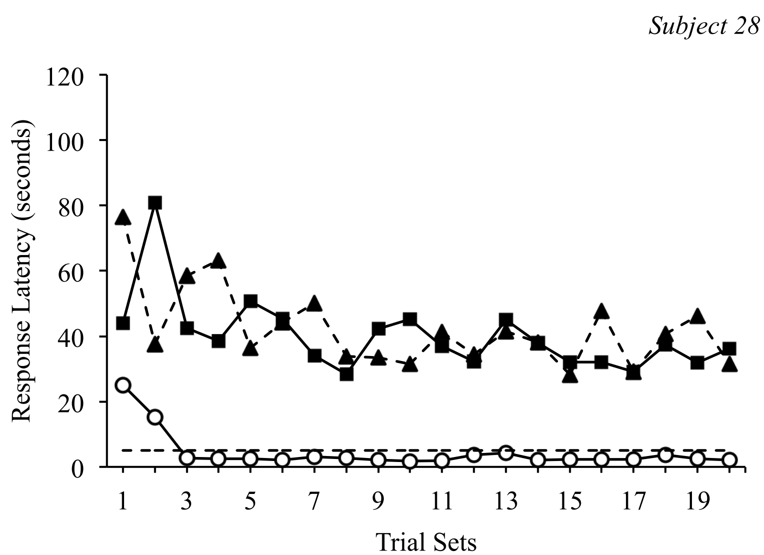


FIGURE 35b. Response latency measures across Control, CD and CID conditions.

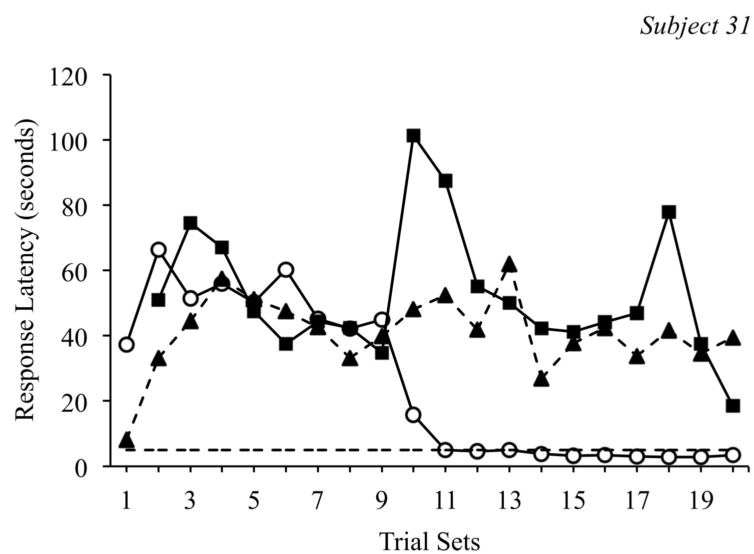


FIGURE 36b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 2: Avoidance Group Subjects

Subject 33

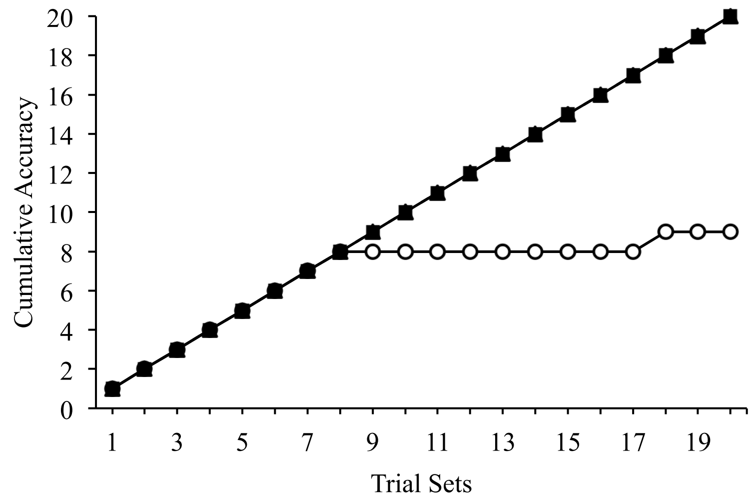


FIGURE 37a. Cumulative accuracy across Control, CD and CID conditions.

Subject 34

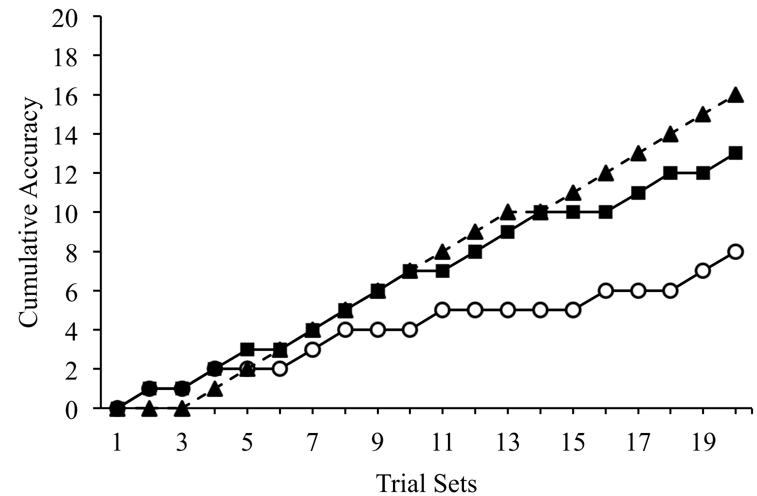


FIGURE 38a. Cumulative accuracy across Control, CD and CID conditions.

Subject 33

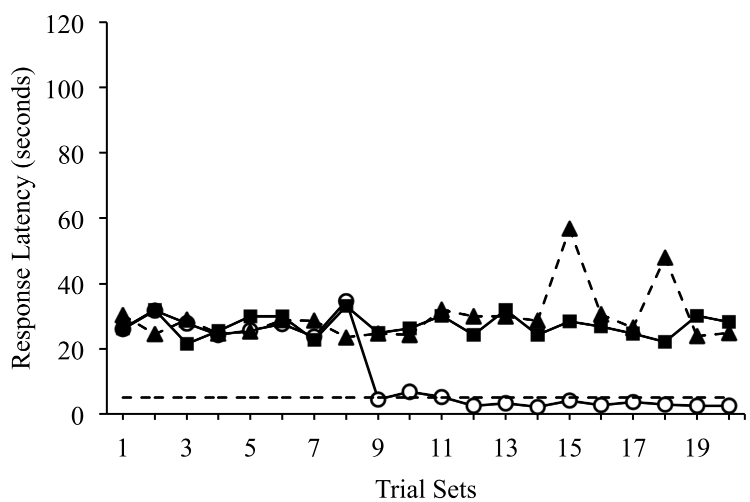


FIGURE 37b. Response latency measures across Control, CD and CID conditions.

Subject 34

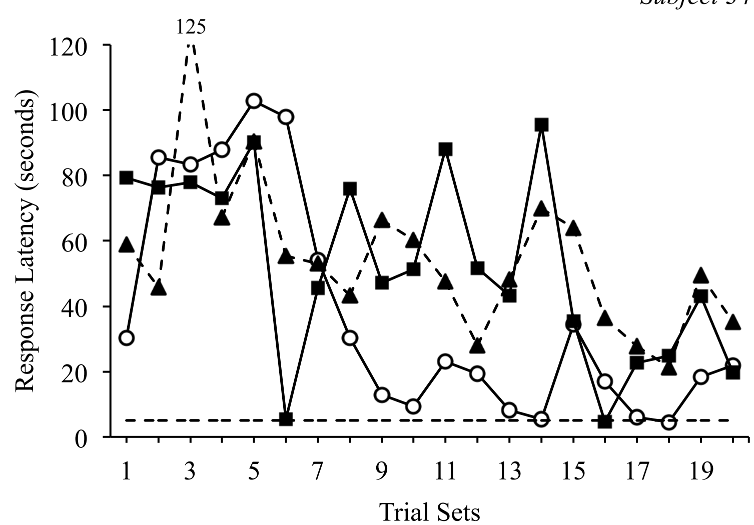


FIGURE 38b. Response latency measures across Control, CD and CID conditions.

Control — Contingent Delay — Contingent Interactive Delay — 5-seconds

Experiment 2: Avoidance Group Subjects

Subject 37

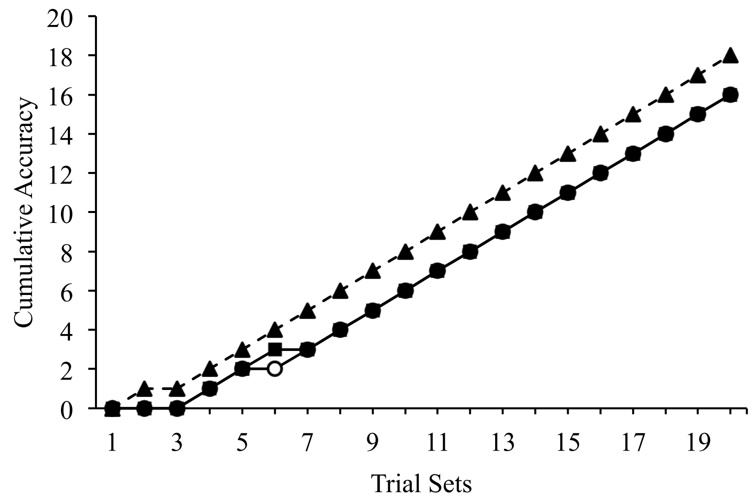


FIGURE 39a. Cumulative accuracy across Control, CD and CID conditions.

Subject 40

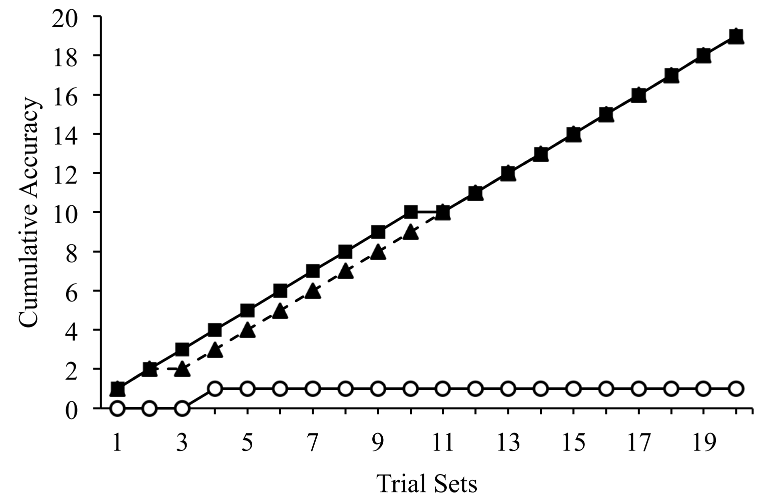


FIGURE 40a. Cumulative accuracy across Control, CD and CID conditions.

Subject 37

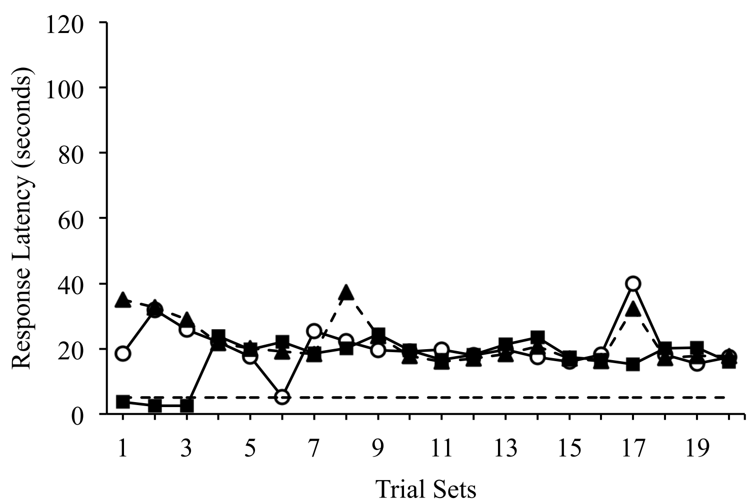


FIGURE 39b. Response latency measures across Control, CD and CID conditions.

Subject 40

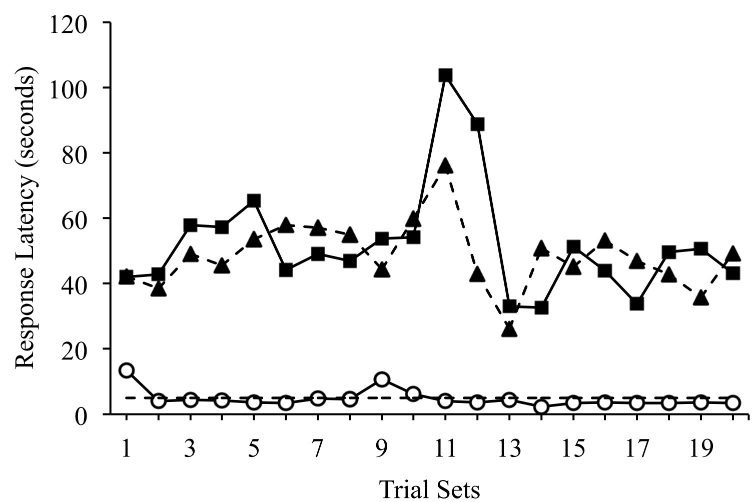


FIGURE 40b. Response latency measures across Control, CD and CID conditions.

Control Contingent Delay Contingent Interactive Delay 5-seconds

Experiment 2: Avoidance Group Subjects

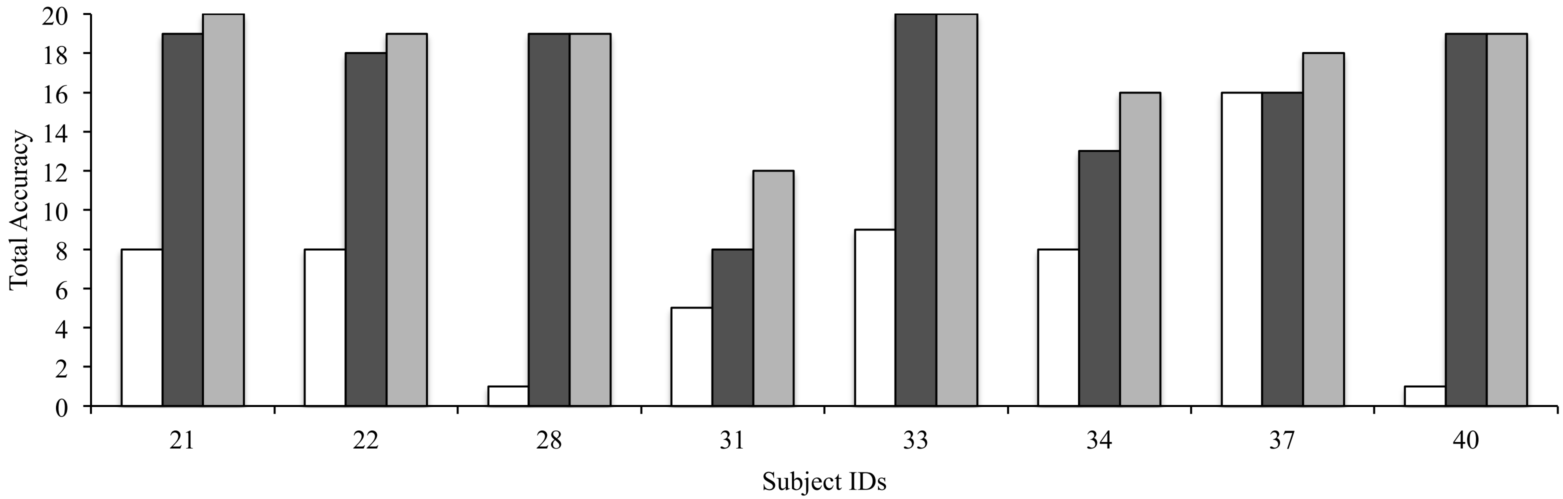


FIGURE 41a. Total accuracy of responding for all Subjects in the Avoidance Group across Control, CD and CID conditions.

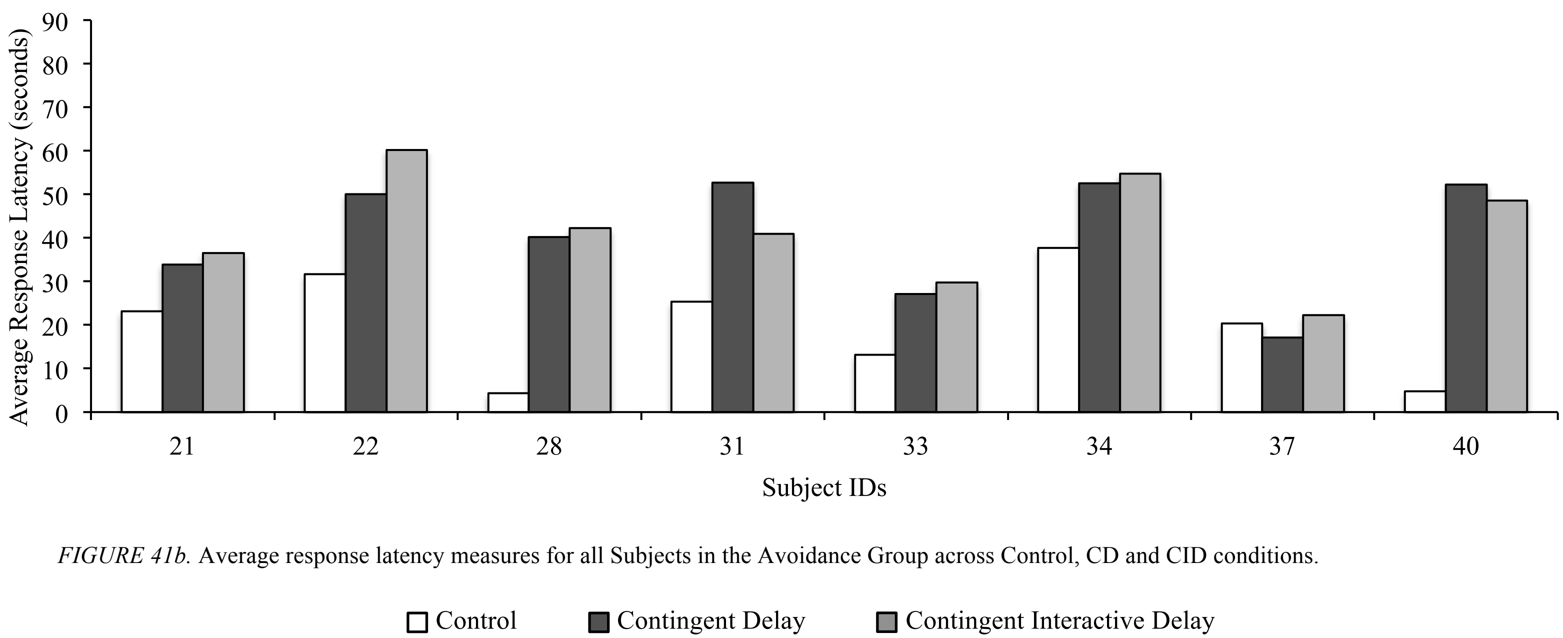


FIGURE 41b. Average response latency measures for all Subjects in the Avoidance Group across Control, CD and CID conditions.

Experiment 2: Endurance Group Subjects

Subject 24

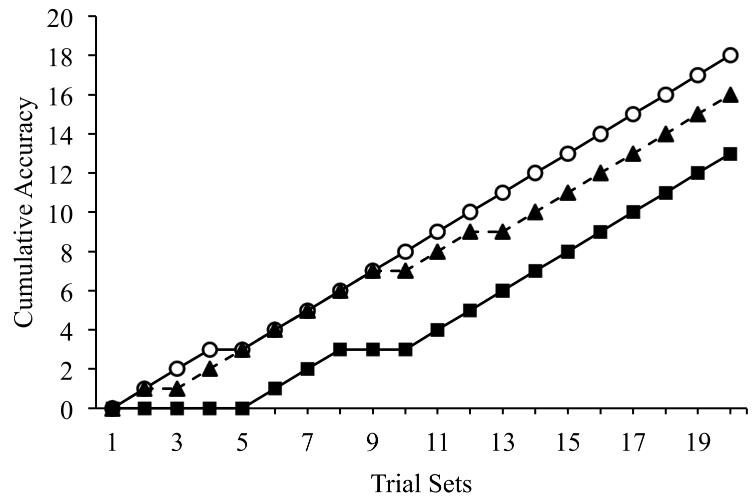


FIGURE 42a. Cumulative accuracy across Control, CD and CID conditions.

Subject 25

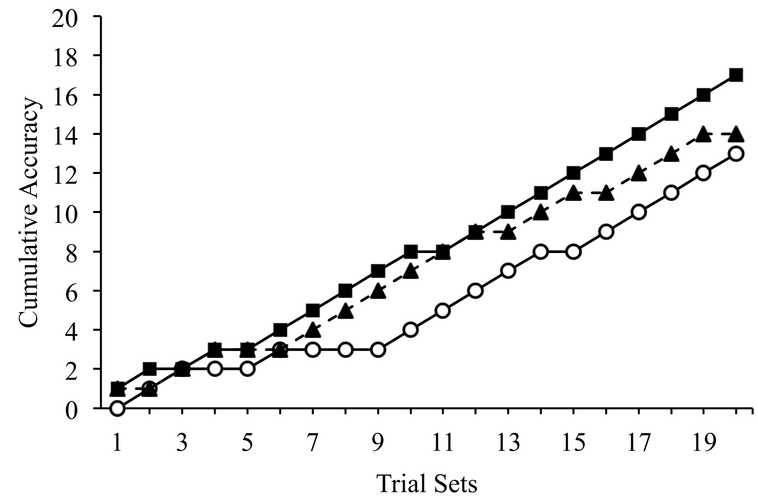


FIGURE 43a. Cumulative accuracy across Control, CD and CID conditions.

Subject 24

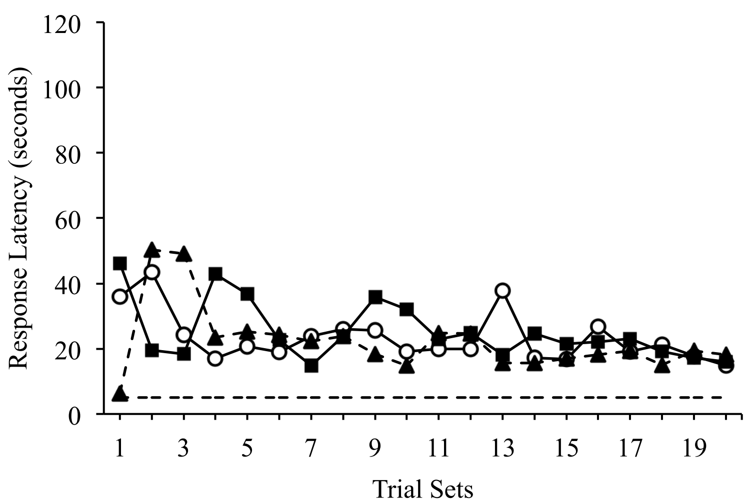


FIGURE 42b. Response latency measures across Control, CD and CID conditions.

Subject 25

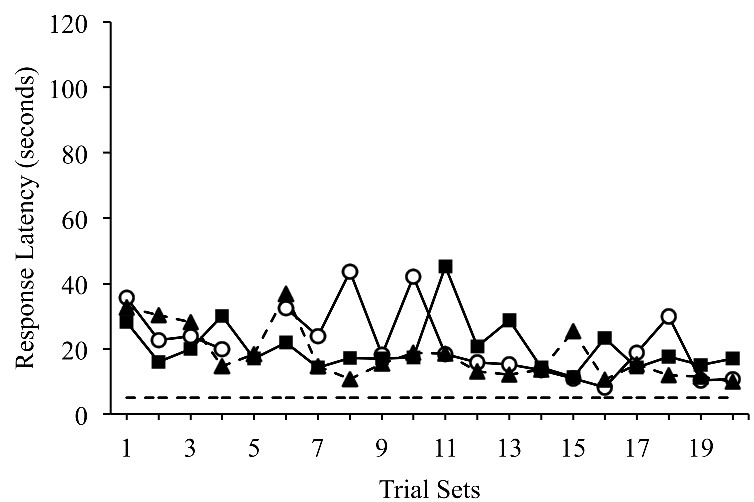


FIGURE 43b. Response latency measures across Control, CD and CID conditions.

Control　Contingent Delay　Contingent Interactive Delay　5-seconds

Experiment 2: Endurance Group Subjects

Subject 35

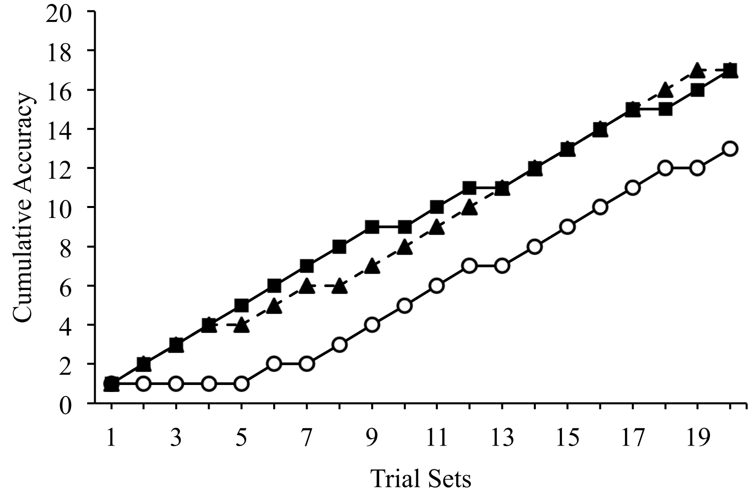


FIGURE 44a. Cumulative accuracy across Control, CD and CID conditions.

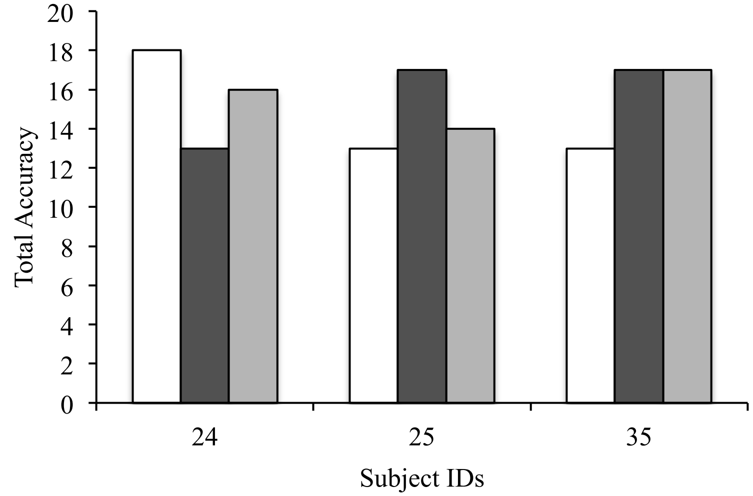


FIGURE 45a. Total accuracy of responding for all Subjects in the Endurance Group across Control, CD and CID conditions.

Subject 35

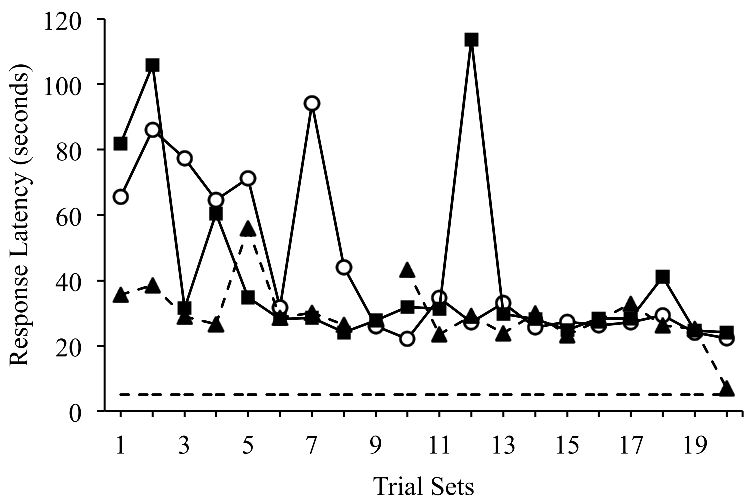


FIGURE 44b. Response latency measures across Control, CD and CID conditions.

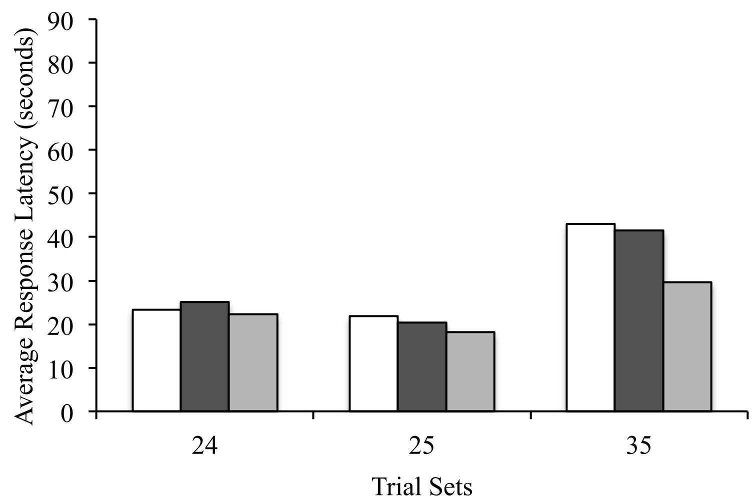


FIGURE 45b. Average response latency measures for all Subjects in the Endurance Group across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Control   Contingent Delay   Contingent Interactive Delay

Experiment 2: Inattentive Group Subject

Subject 38

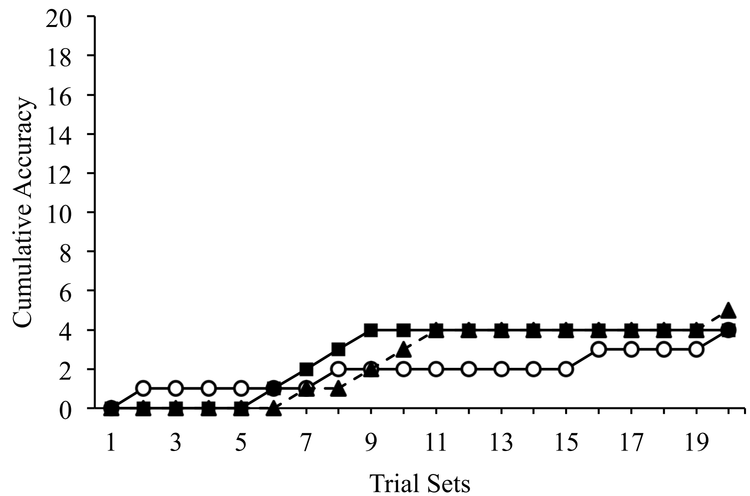


FIGURE 46a. Cumulative accuracy across Control, CD and CID conditions.

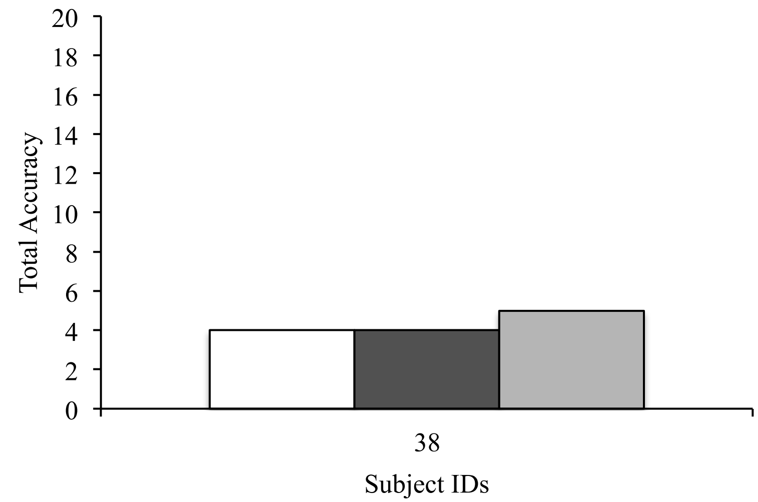


FIGURE 47a. Total accuracy of responding for all Subjects in the Inattentive Group across Control, CD and CID conditions.

Subject 38

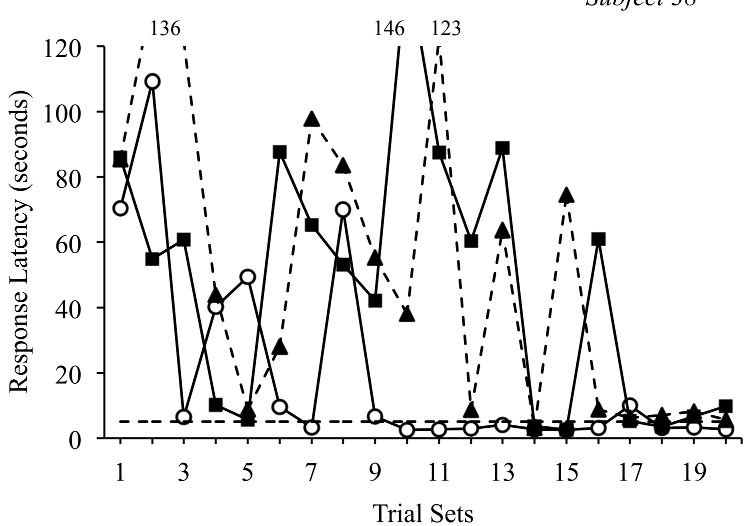


FIGURE 46b. Response latency measures across Control, CD and CID conditions.

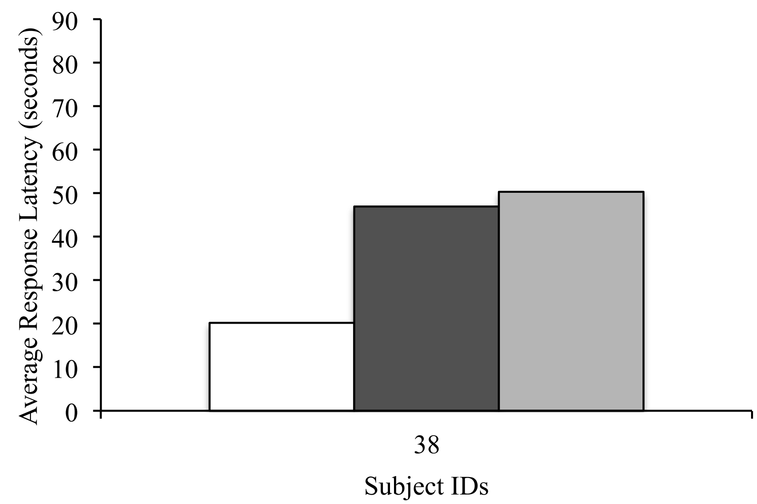


FIGURE 47b. Average response latency measures for all Subjects in the Inattentive Group across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Control   Contingent Delay   Contingent Interactive Delay

Experiment 3: Proficiency Group Subjects

Subject 41

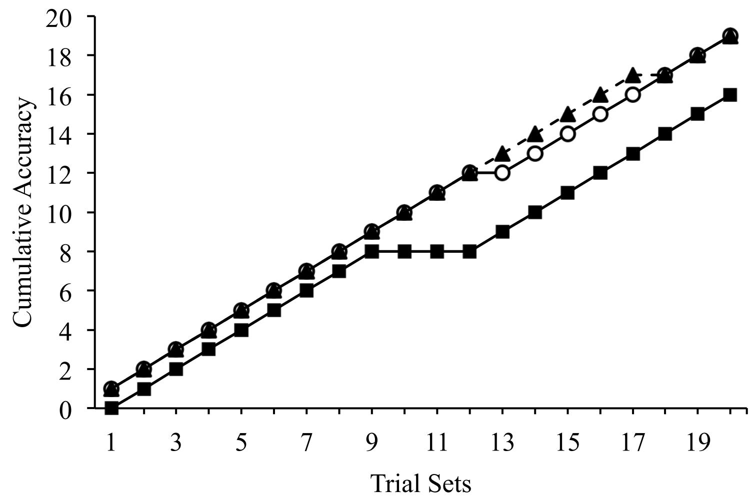


FIGURE 48a. Cumulative accuracy across Control, CD and CID conditions.

Subject 42

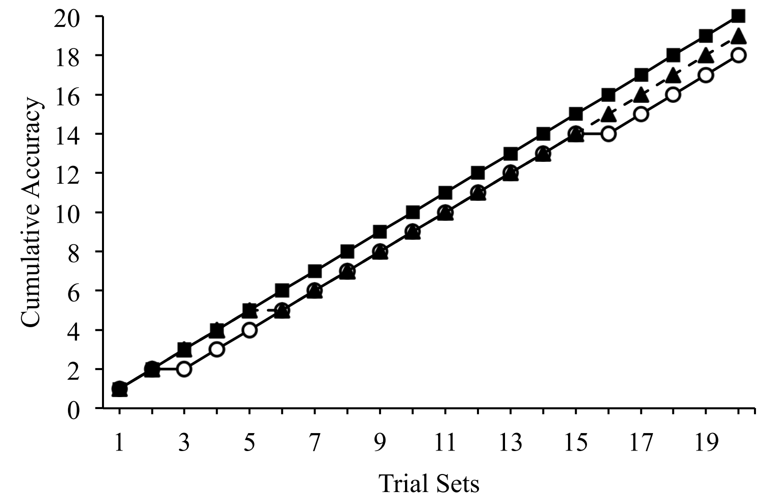


FIGURE 49a. Cumulative accuracy across Control, CD and CID conditions.

Subject 41

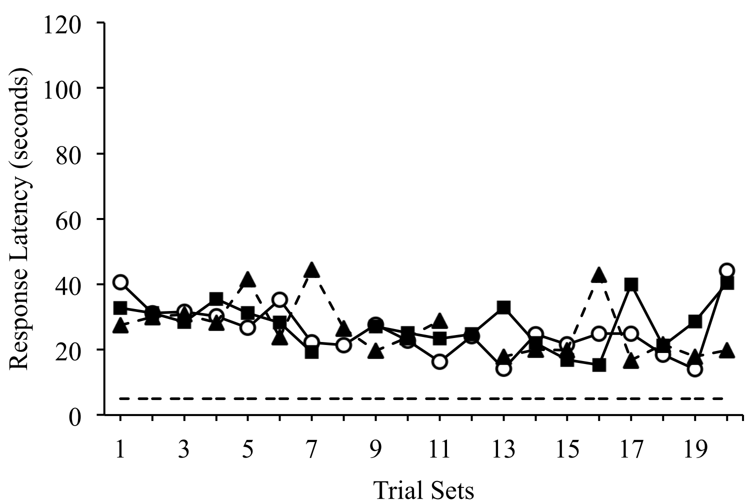


FIGURE 48b. Response latency measures across Control, CD and CID conditions.

Subject 42

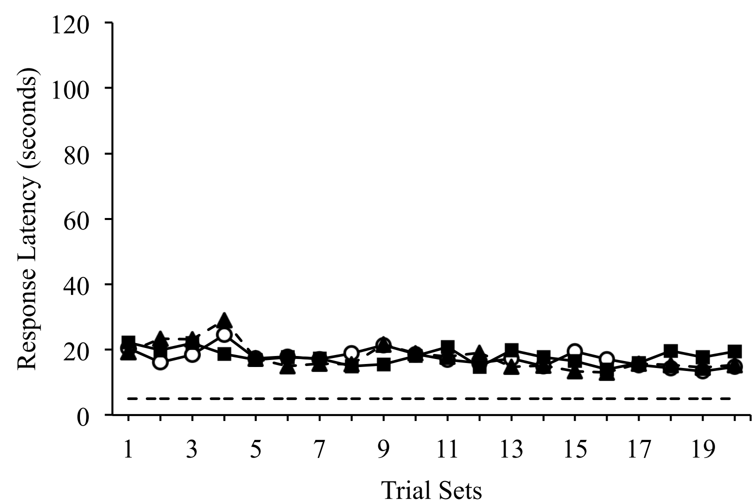


FIGURE 49b. Response latency measures across Control, CD and CID conditions.

Control Contingent Delay Contingent Interactive Delay 5-seconds

Experiment 3: Proficiency Group Subjects

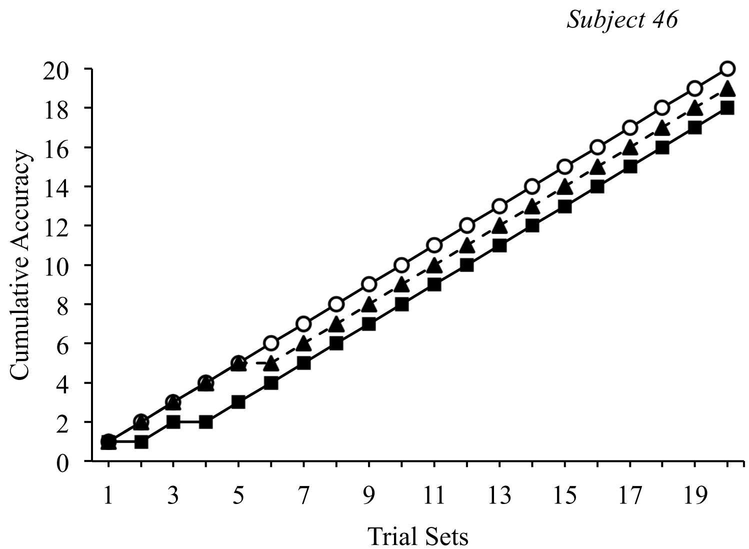


FIGURE 50a. Cumulative accuracy across Control, CD and CID conditions.

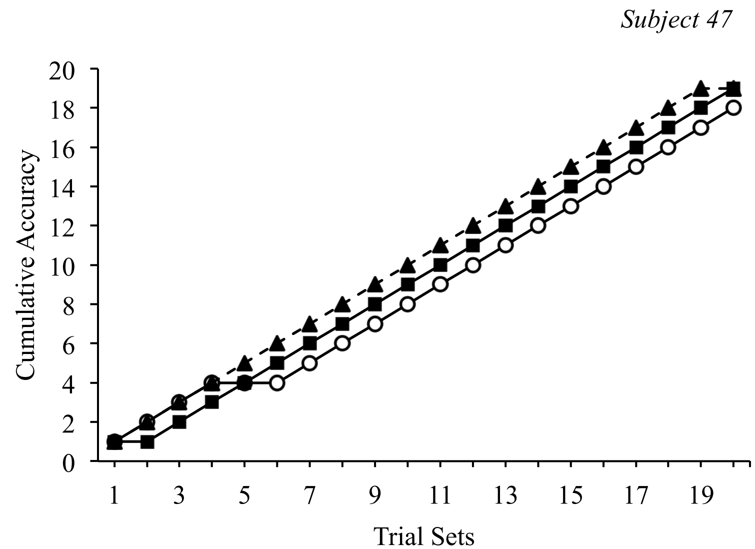


FIGURE 51a. Cumulative accuracy across Control, CD and CID conditions.

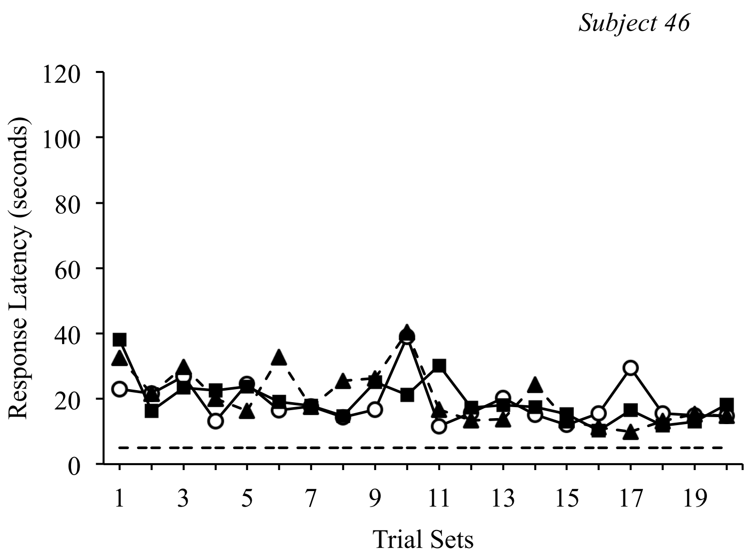


FIGURE 50b. Response latency measures across Control, CD and CID conditions.

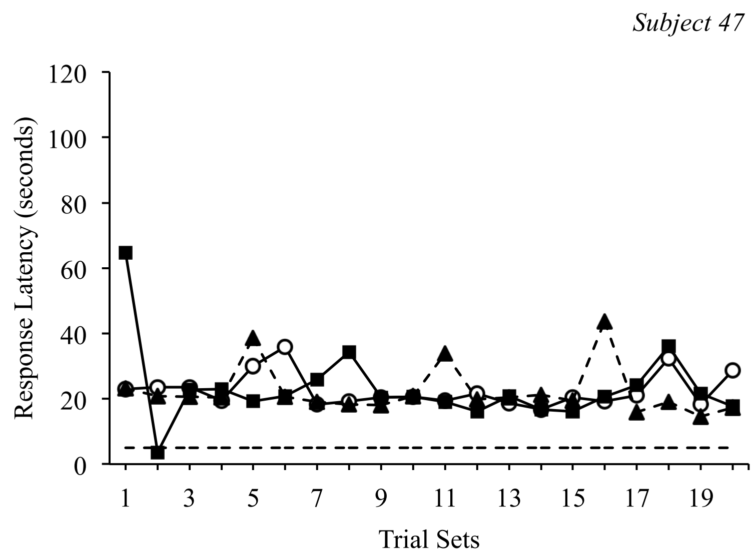


FIGURE 51b. Response latency measures across Control, CD and CID conditions.

Control Contingent Delay Contingent Interactive Delay 5-seconds

Experiment 3: Proficiency Group Subjects

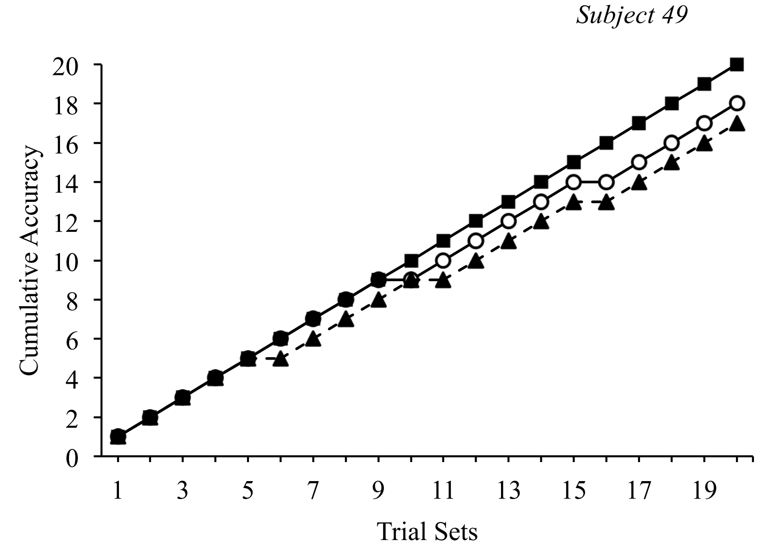


FIGURE 52a. Cumulative accuracy across Control, CD and CID conditions.

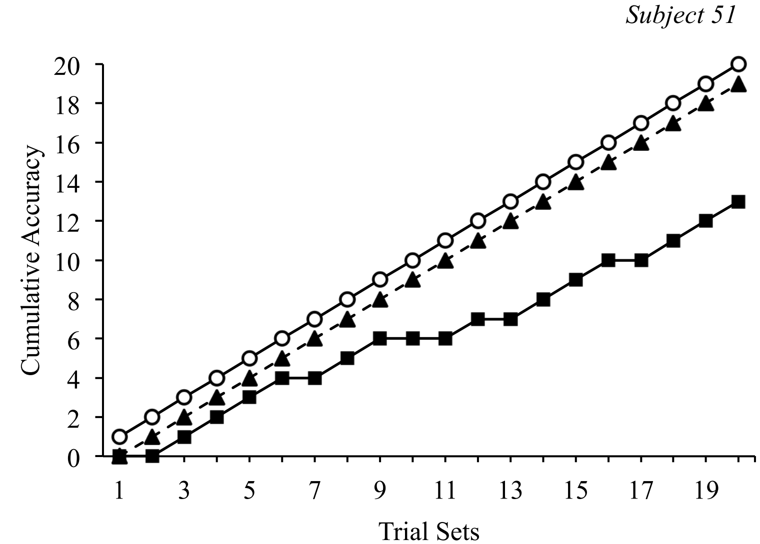


FIGURE 53a. Cumulative accuracy across Control, CD and CID conditions.

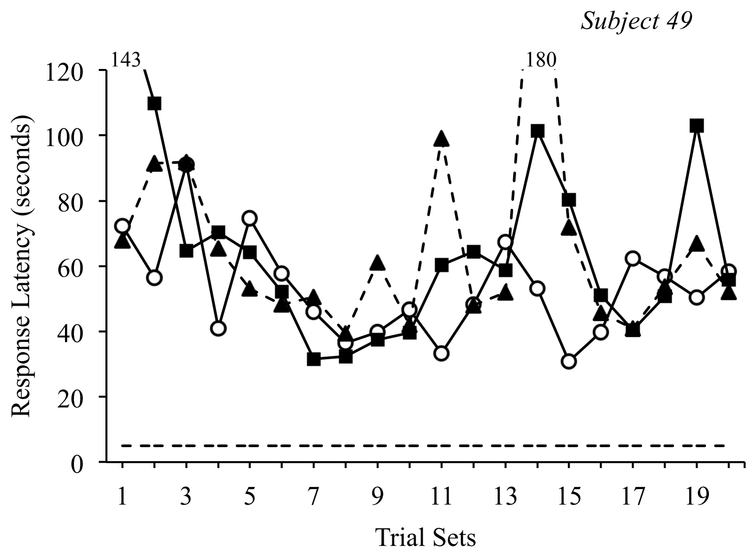


FIGURE 52b. Response latency measures across Control, CD and CID conditions.

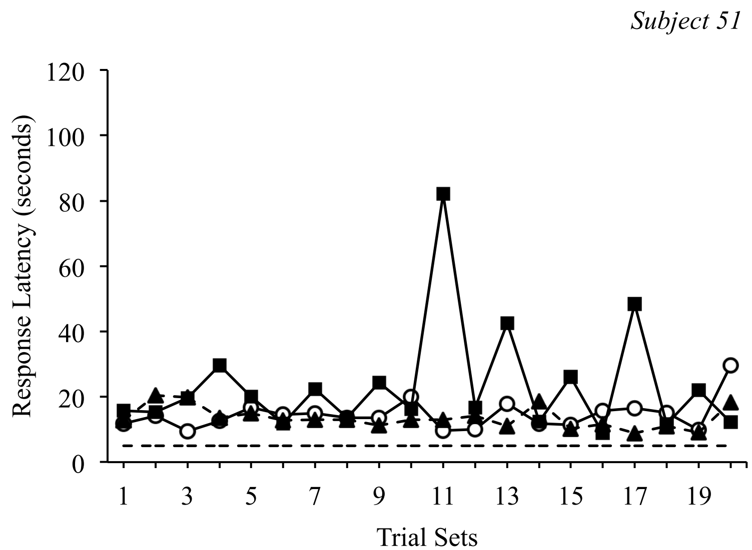


FIGURE 53b. Response latency measures across Control, CD and CID conditions.

Control Contingent Delay Contingent Interactive Delay 5-seconds

Experiment 3: Proficiency Group Subjects

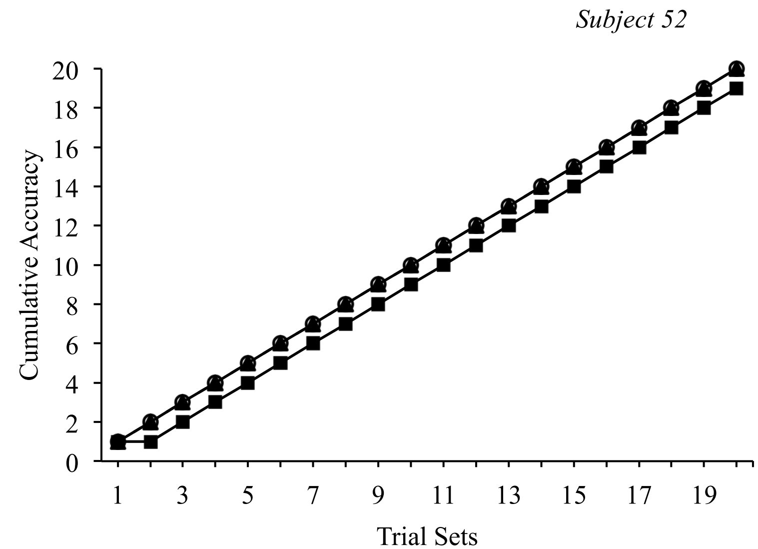


FIGURE 54a. Cumulative accuracy across Control, CD and CID conditions.

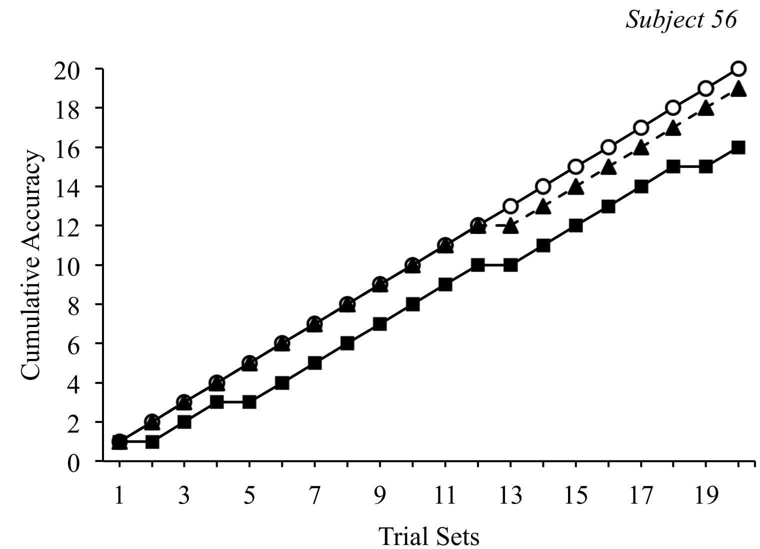


FIGURE 55a. Cumulative accuracy across Control, CD and CID conditions.

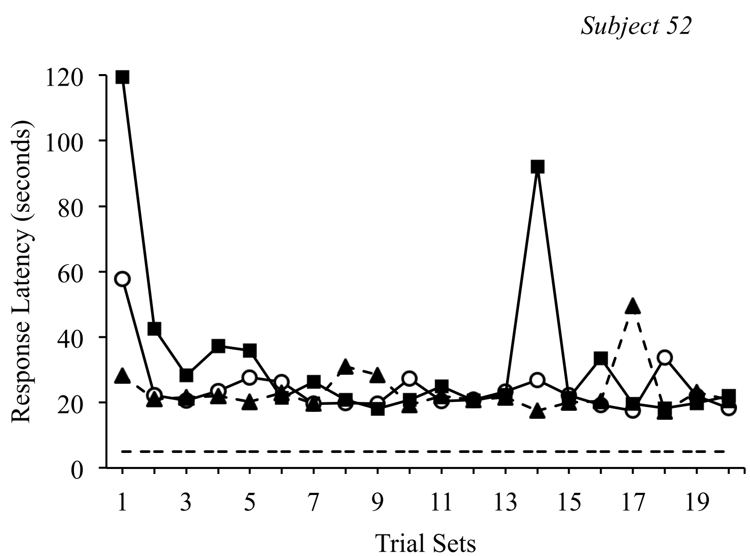


FIGURE 54b. Response latency measures across Control, CD and CID conditions.

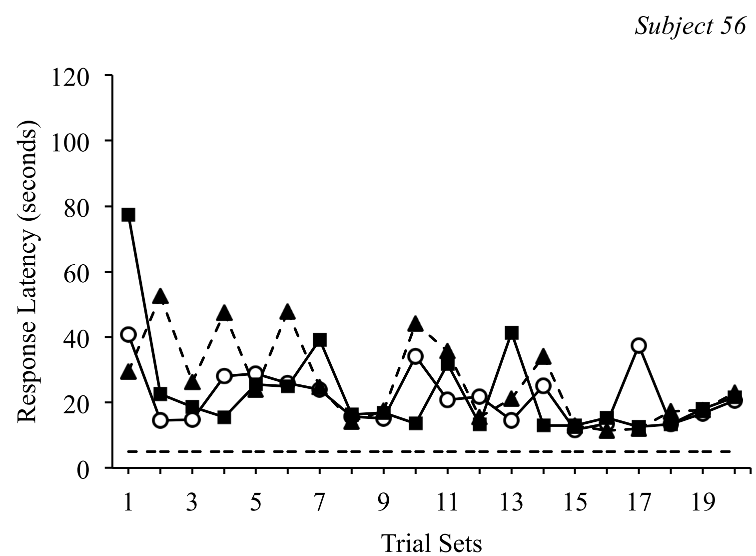


FIGURE 55b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 3: Proficiency Group Subjects

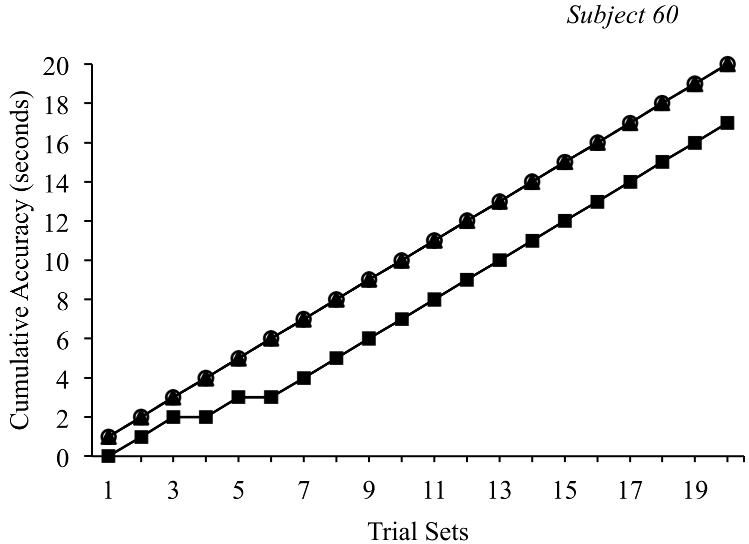


FIGURE 56a. Cumulative accuracy across Control, CD and CID conditions.

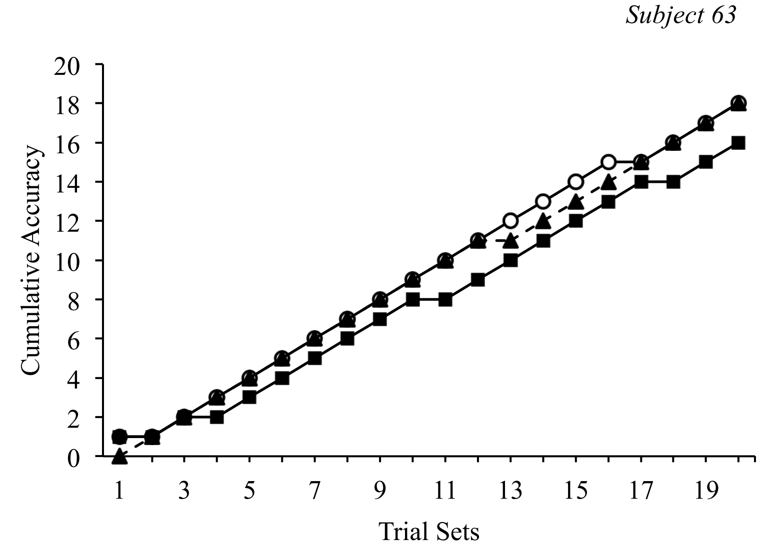


FIGURE 57a. Cumulative accuracy across Control, CD and CID conditions.

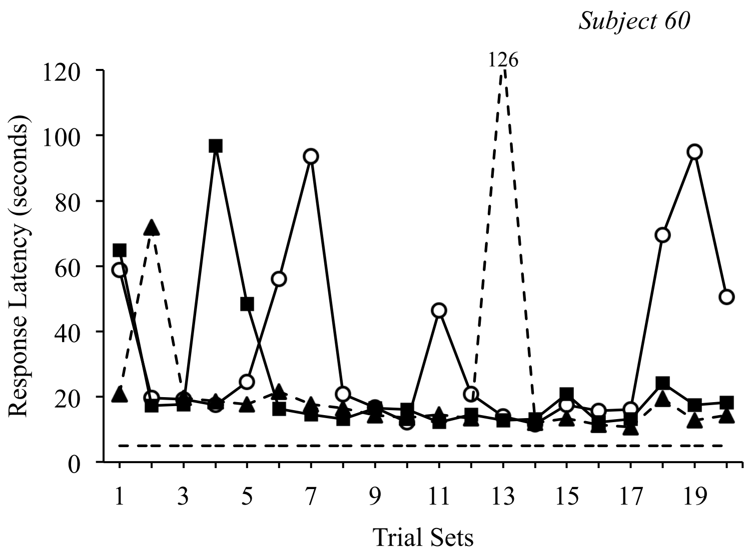


FIGURE 56b. Response latency measures across Control, CD and CID conditions.

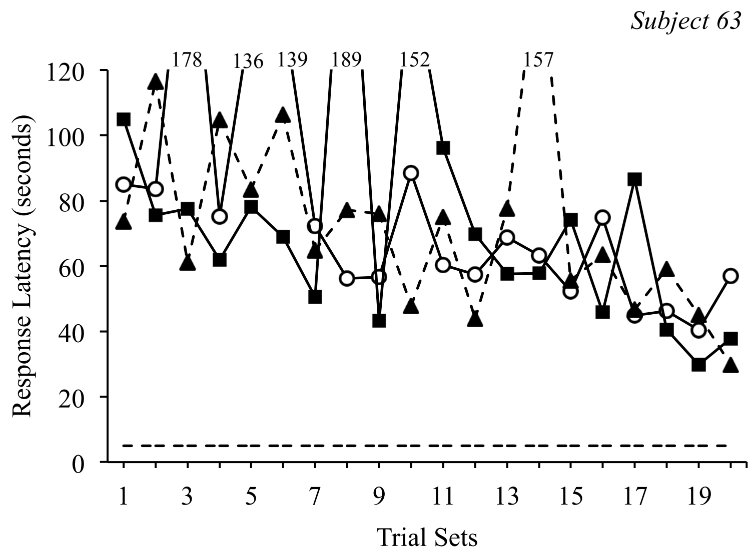
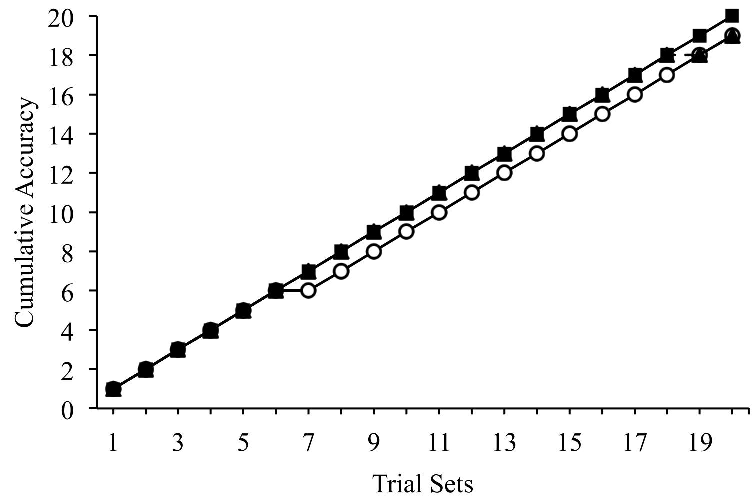


FIGURE 57b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 3: Proficiency Group Subjects

Subject 66



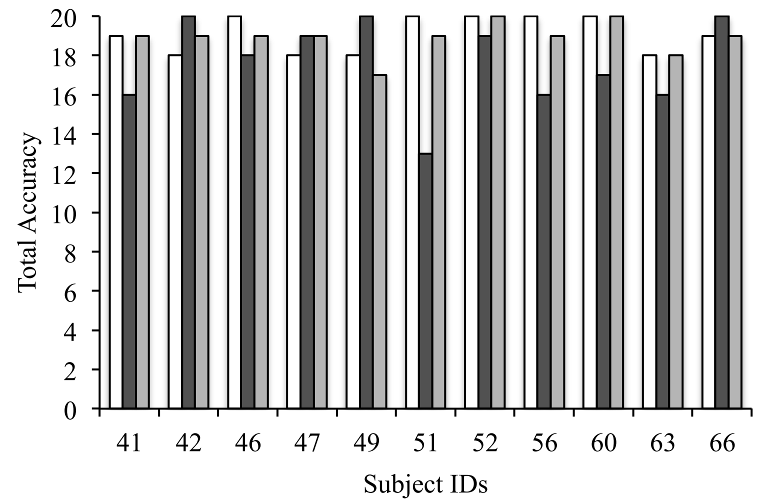
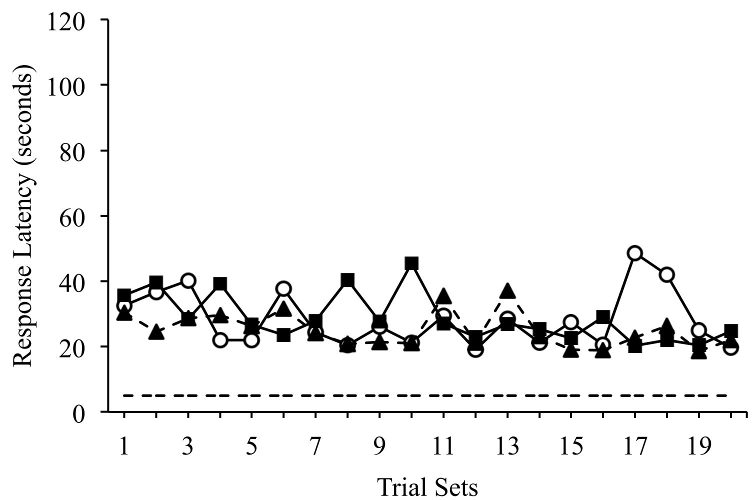


FIGURE 58a. Cumulative accuracy across Control, CD and CID conditions.

FIGURE 59a. Total accuracy of responding for all Subjects in the Proficiency Group across Control, CD and CID conditions.

Subject 66



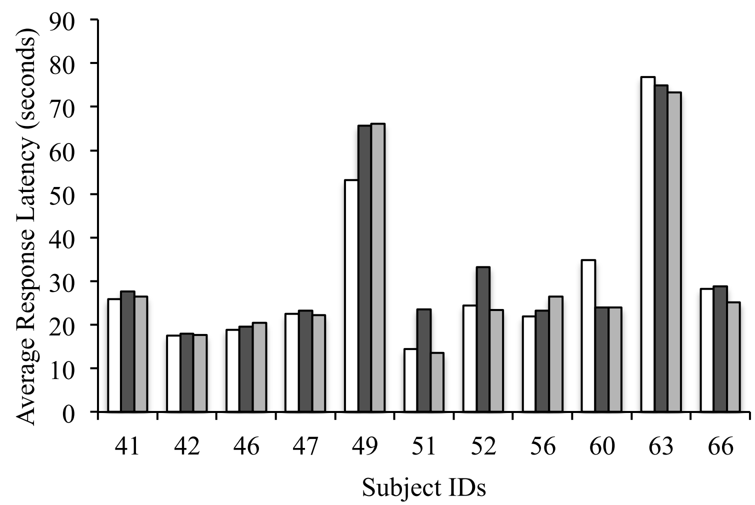


FIGURE 58b. Response latency measures across Control, CD and CID conditions.

Control Contingent Delay Contingent Interactive Delay 5-seconds

FIGURE 59b. Average response latency measures for all Subjects in the Proficiency Group across Control, CD and CID conditions.

Control Contingent Delay Contingent Interactive Delay

Experiment 3: Avoidance Group Subjects

Subject 43

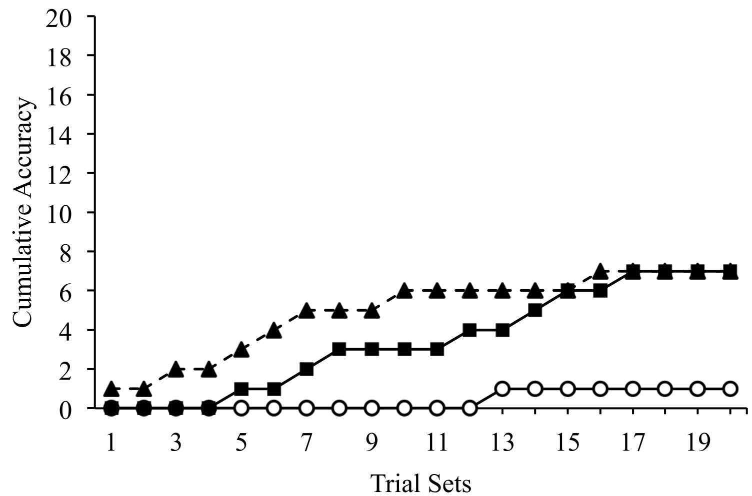


FIGURE 60a. Cumulative accuracy across Control, CD and CID conditions.

Subject 53

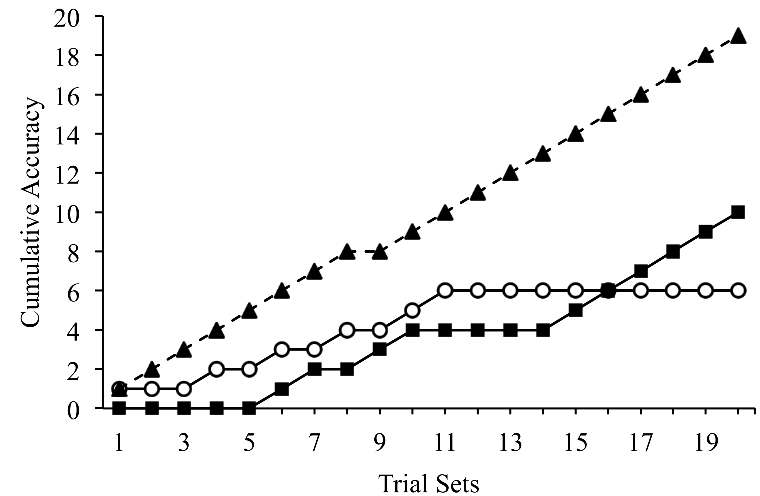


FIGURE 61a. Cumulative accuracy across Control, CD and CID conditions.

Subject 43

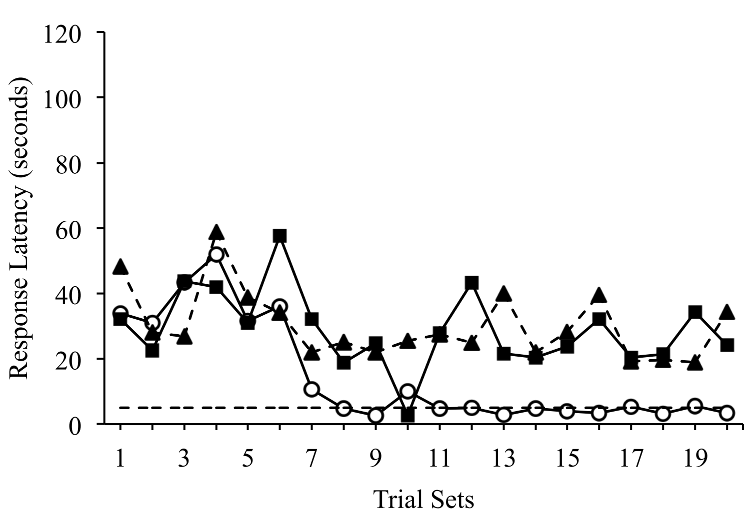


FIGURE 60b. Response latency measures across Control, CD and CID conditions.

Subject 53

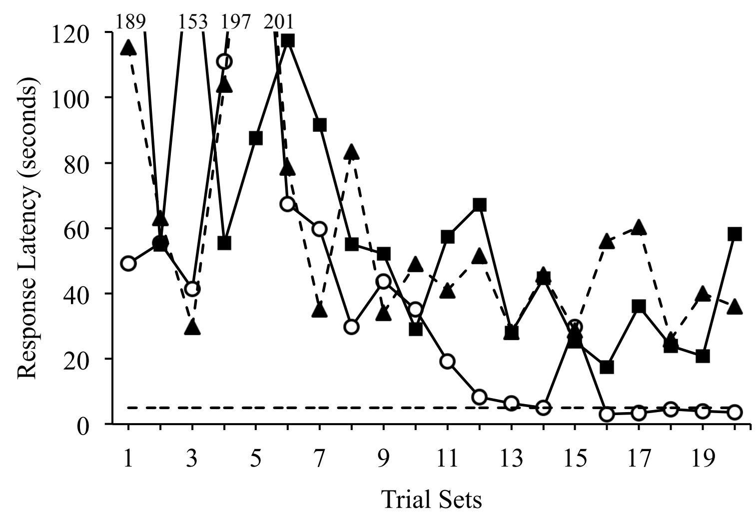


FIGURE 61b. Response latency measures across Control, CD and CID conditions.

Control — Contingent Delay — Contingent Interactive Delay — 5-seconds

Experiment 3: Avoidance Group Subjects

Subject 55

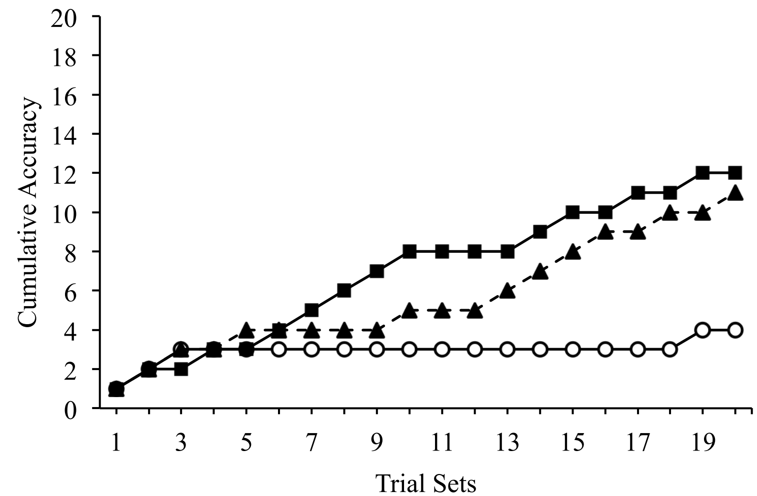


FIGURE 62a. Cumulative accuracy across Control, CD and CID conditions.

Subject 58

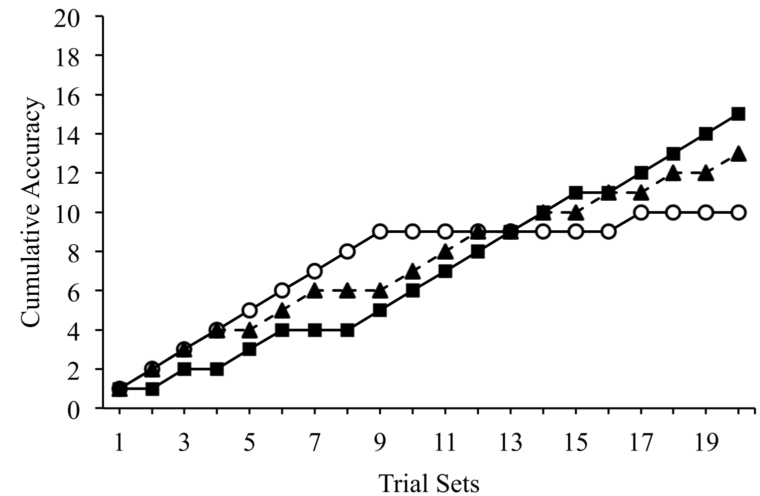


FIGURE 63a. Cumulative accuracy across Control, CD and CID conditions.

Subject 55

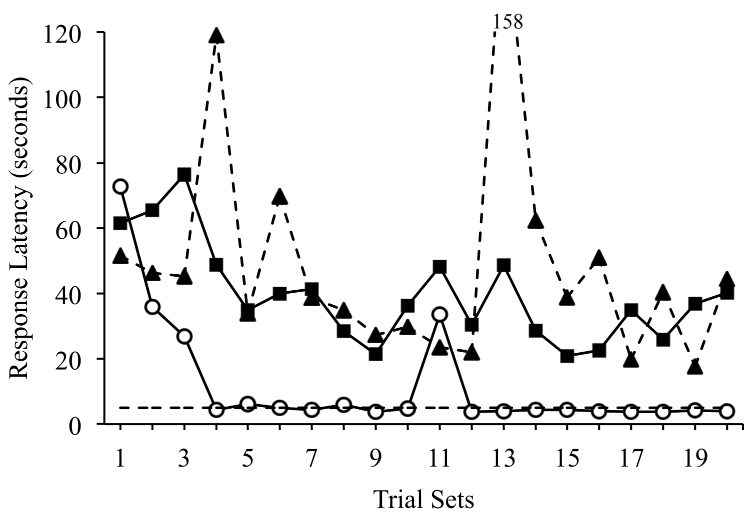


FIGURE 62b. Response latency measures across Control, CD and CID conditions.

Subject 58

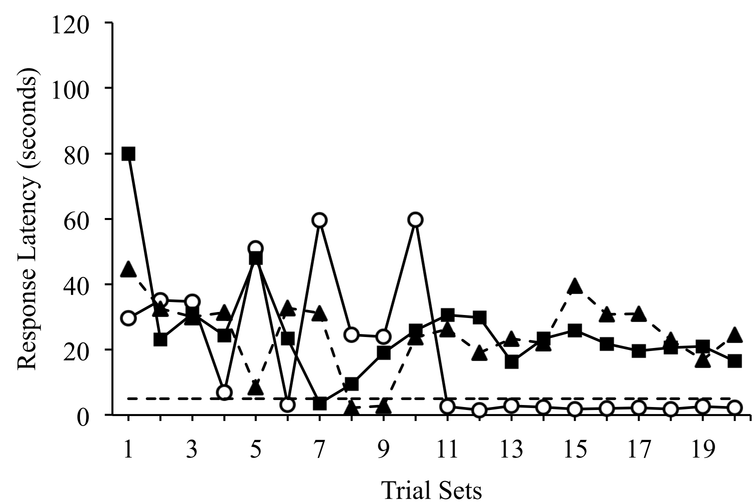


FIGURE 63b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 3: Avoidance Group Subjects

Subject 59

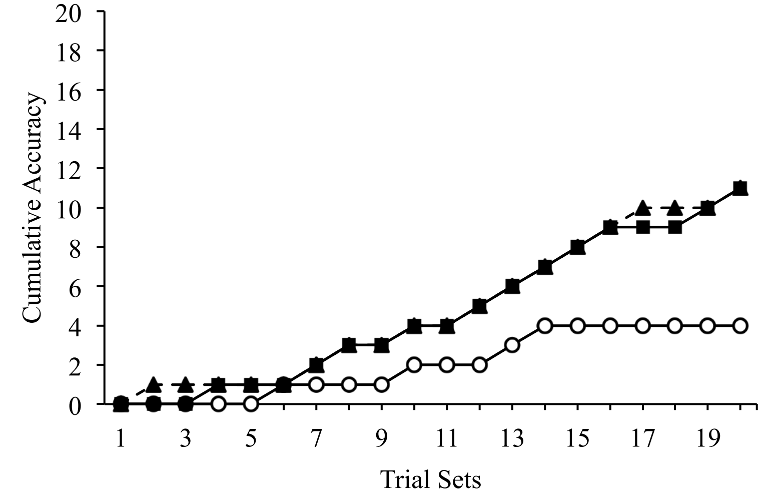


FIGURE 64a. Cumulative accuracy across Control, CD and CID conditions.

Subject 69

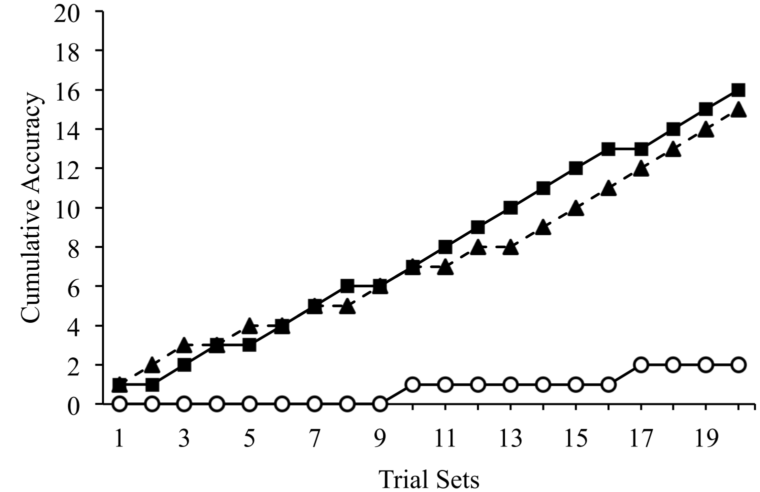


FIGURE 65a. Cumulative accuracy across Control, CD and CID conditions.

Subject 59

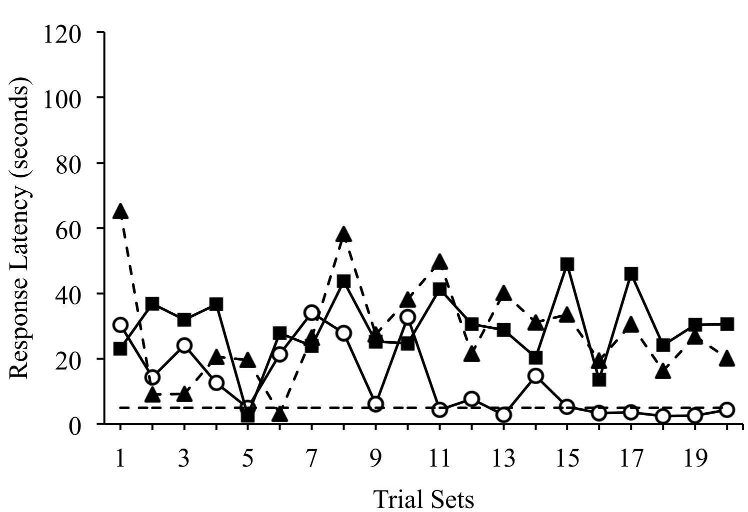


FIGURE 64b. Response latency measures across Control, CD and CID conditions.

Subject 69

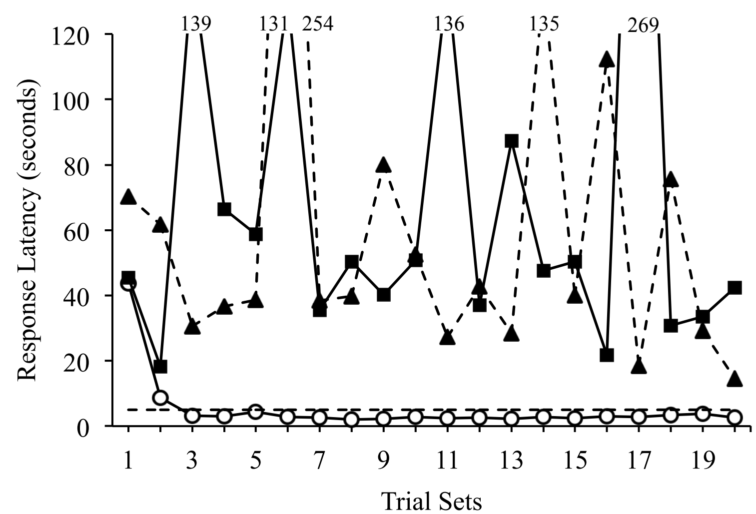


FIGURE 65b. Response latency measures across Control, CD and CID conditions.

Control — Contingent Delay — Contingent Interactive Delay — 5-seconds

Experiment 3: Avoidance Group Subjects

Subject 70

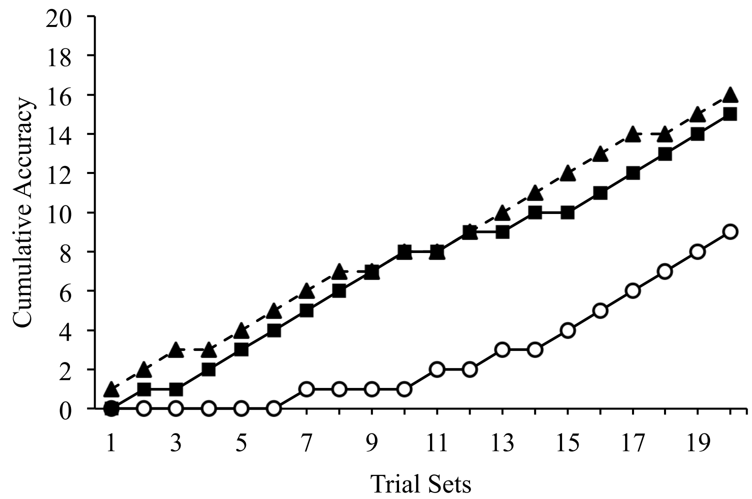


FIGURE 66a. Cumulative accuracy across Control, CD and CID conditions.

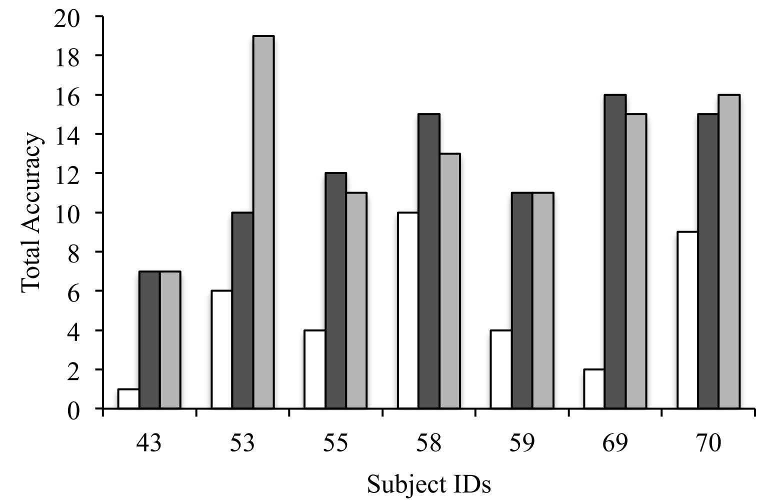


FIGURE 67a. Total accuracy of responding for all Subjects in the Avoidance Group across Control, CD and CID conditions.

Subject 70

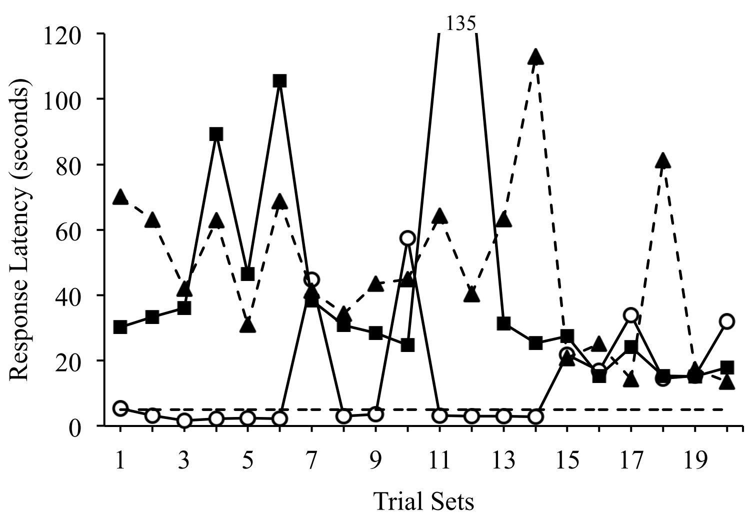


FIGURE 66b. Response latency measures across Control, CD and CID conditions.

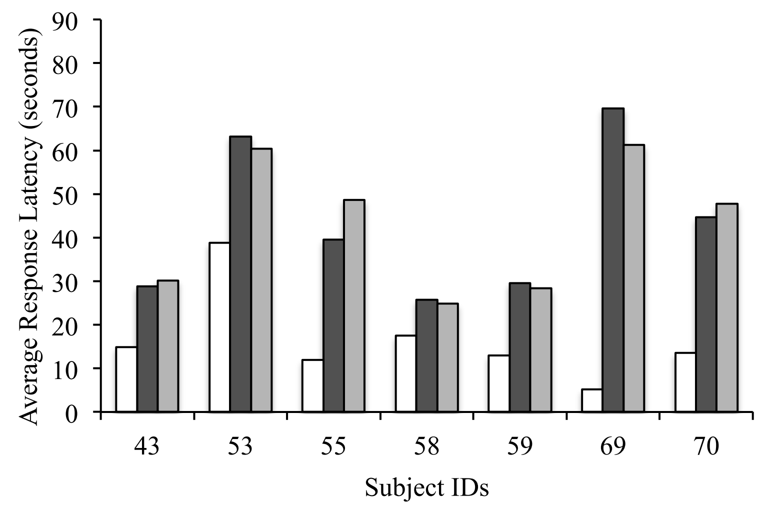


FIGURE 67b. Average response latency measures for all Subjects in the Avoidance Group across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Control   Contingent Delay   Contingent Interactive Delay

Experiment 3: Endurance Group Subjects

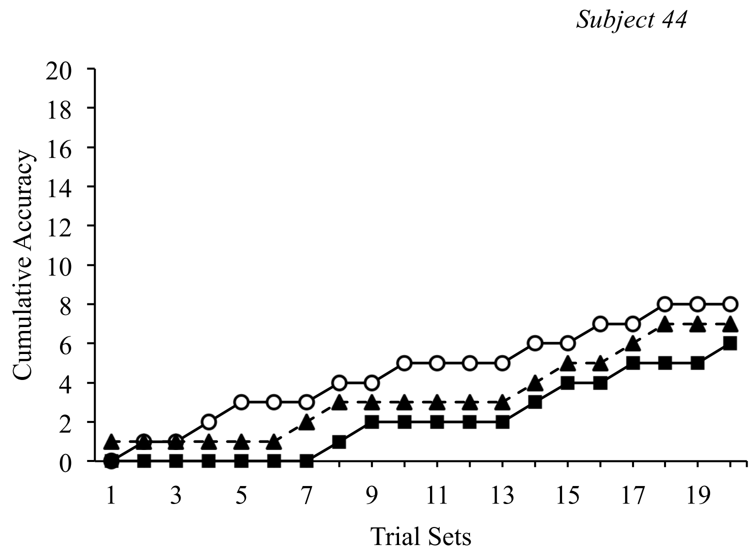


FIGURE 68a. Cumulative accuracy across Control, CD and CID conditions.

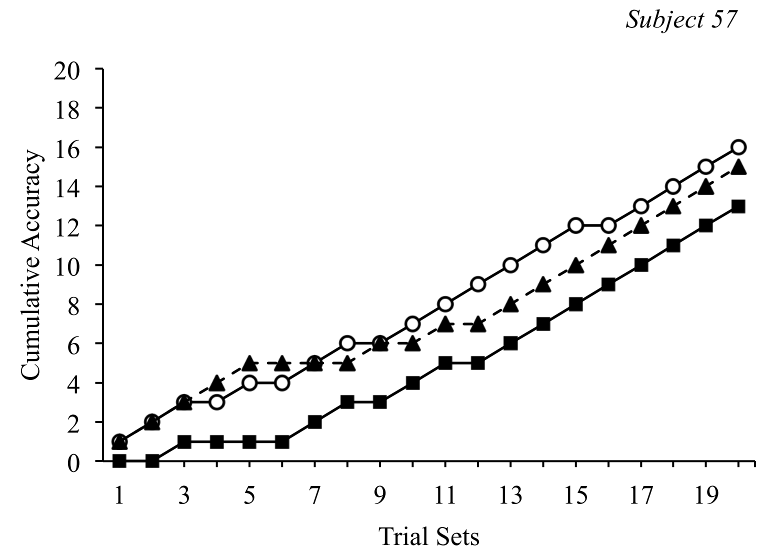


FIGURE 69a. Cumulative accuracy across Control, CD and CID conditions.

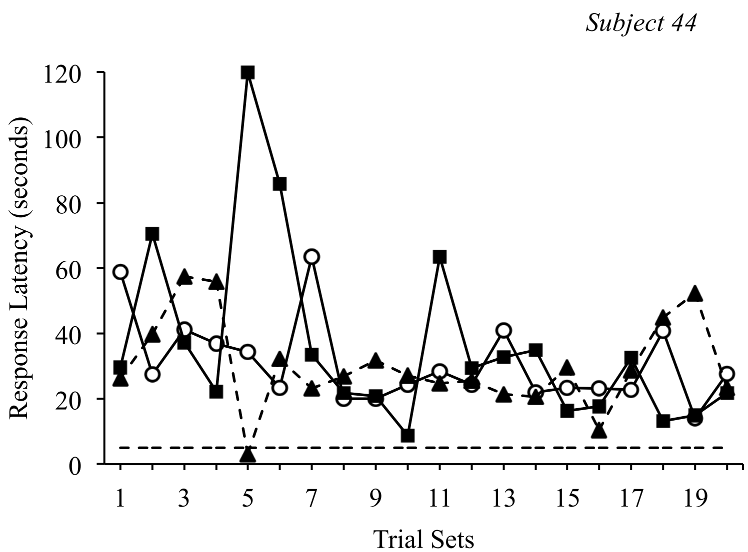


FIGURE 68b. Response latency measures across Control, CD and CID conditions.

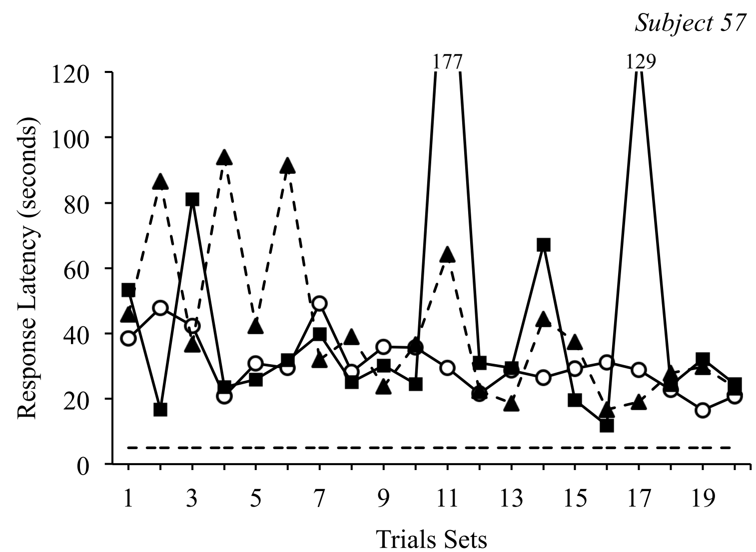


FIGURE 69b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 3: Endurance Group Subjects

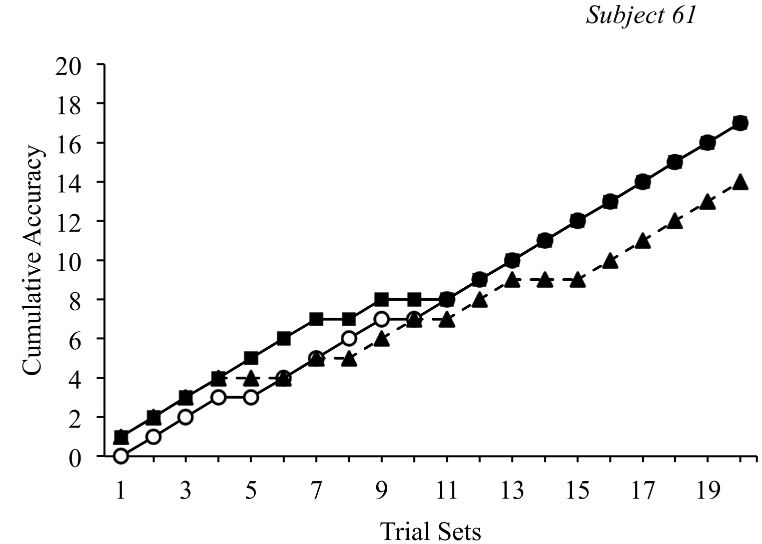


FIGURE 70a. Cumulative accuracy across Control, CD and CID conditions.

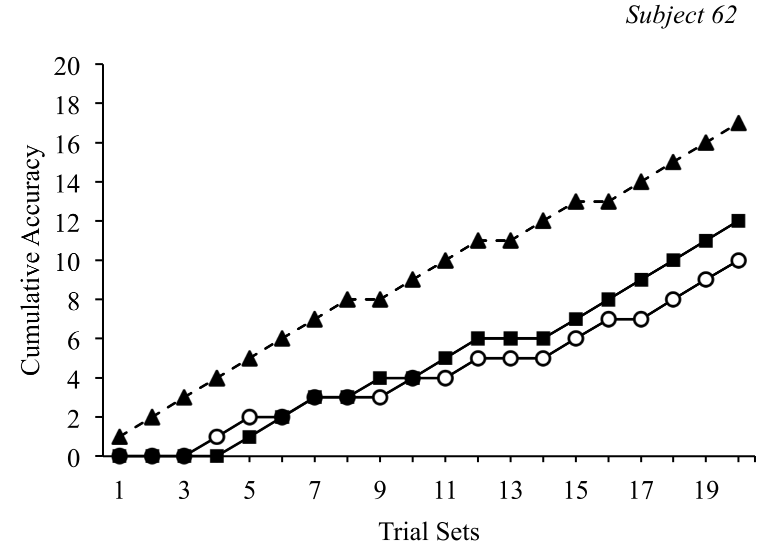


FIGURE 71a. Cumulative accuracy across Control, CD and CID conditions.

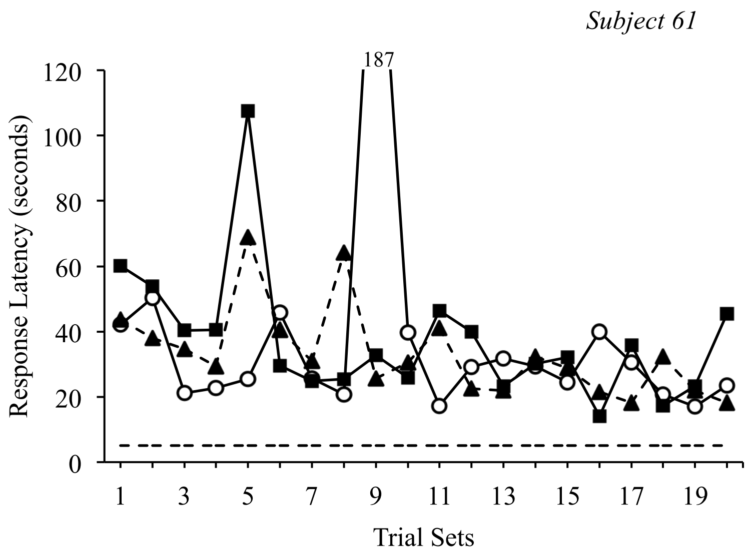


FIGURE 70b. Response latency measures across Control, CD and CID conditions.

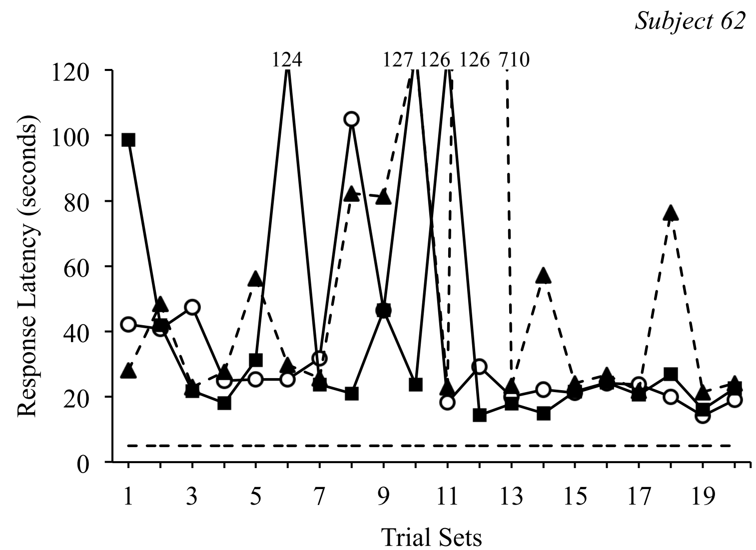


FIGURE 71b. Response latency measures across Control, CD and CID conditions.

Control — Contingent Delay — Contingent Interactive Delay — 5-seconds

Experiment 3: Endurance Group Subjects

Subject 65

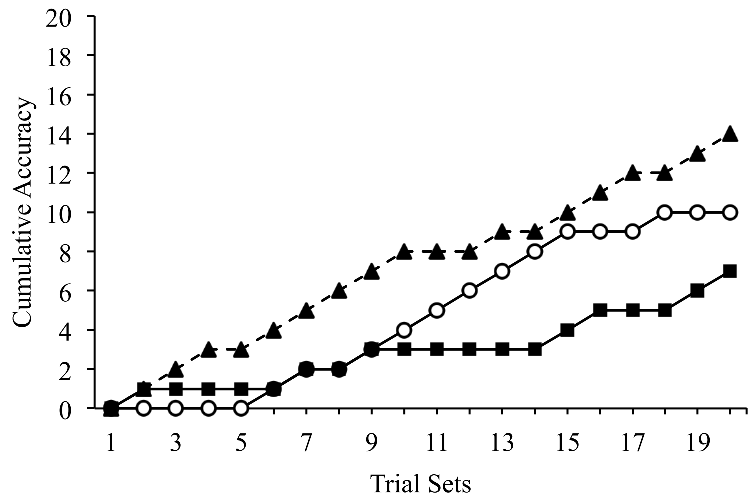


FIGURE 72a. Cumulative accuracy across Control, CD and CID conditions.

Subject 68

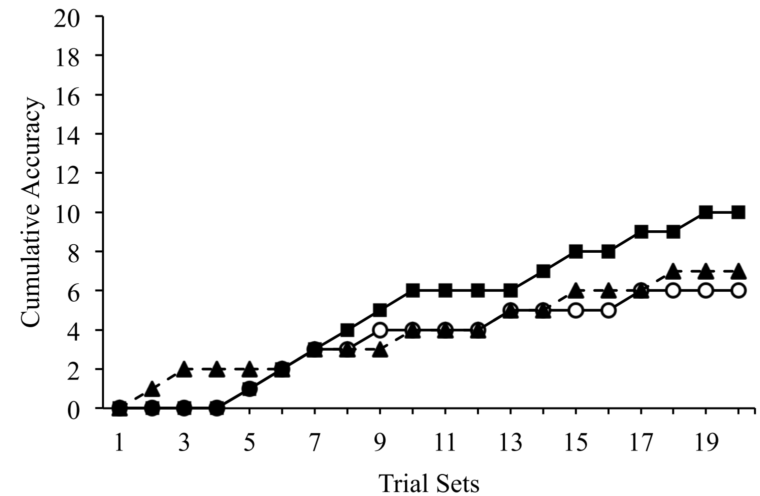


FIGURE 73a. Cumulative accuracy across Control, CD and CID conditions.

Subject 65

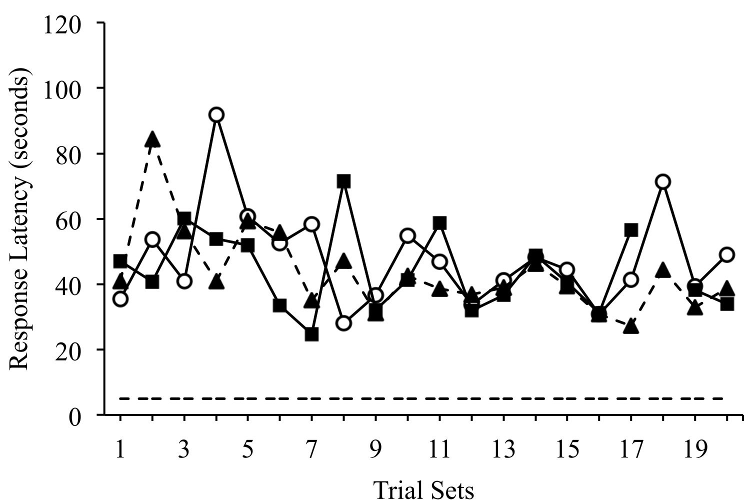


FIGURE 72b. Response latency measures across Control, CD and CID conditions.

Subject 68

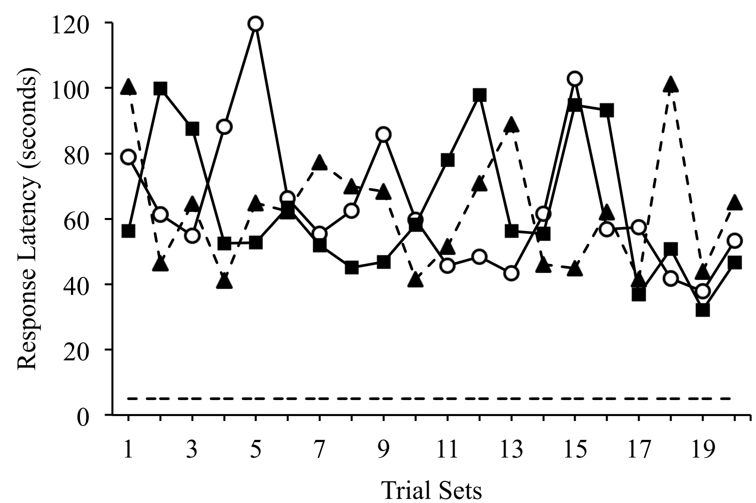


FIGURE 73b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 3: Endurance Group Subjects

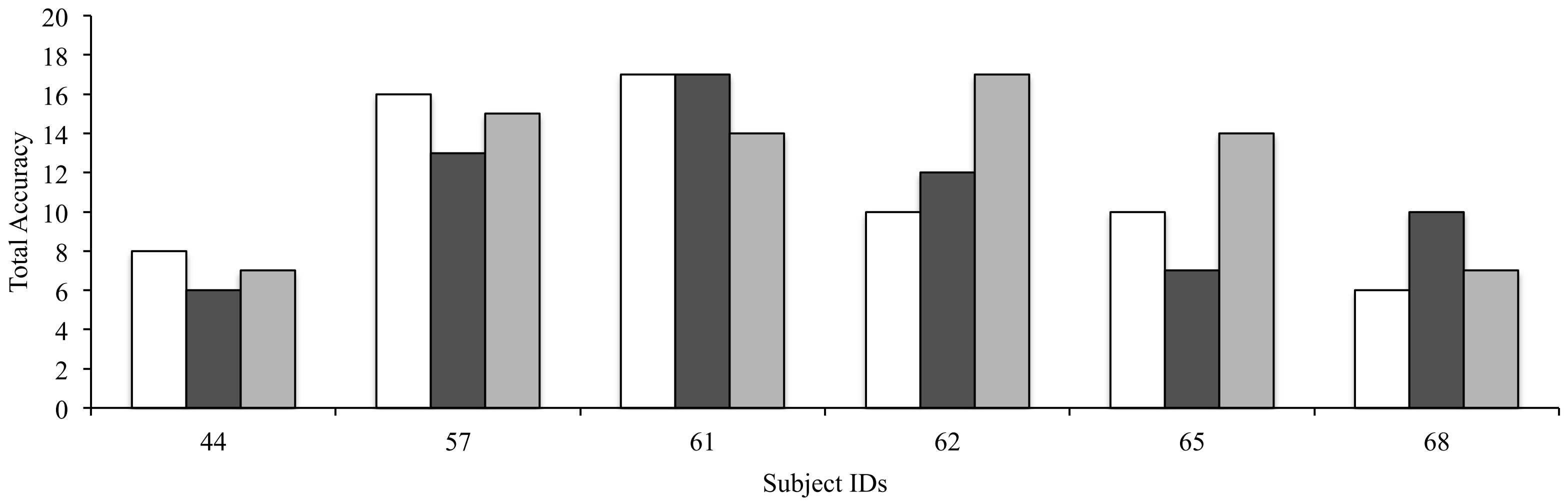


FIGURE 74a. Total accuracy of responding for all Subjects in the Endurance Group across Control, CD and CID conditions.

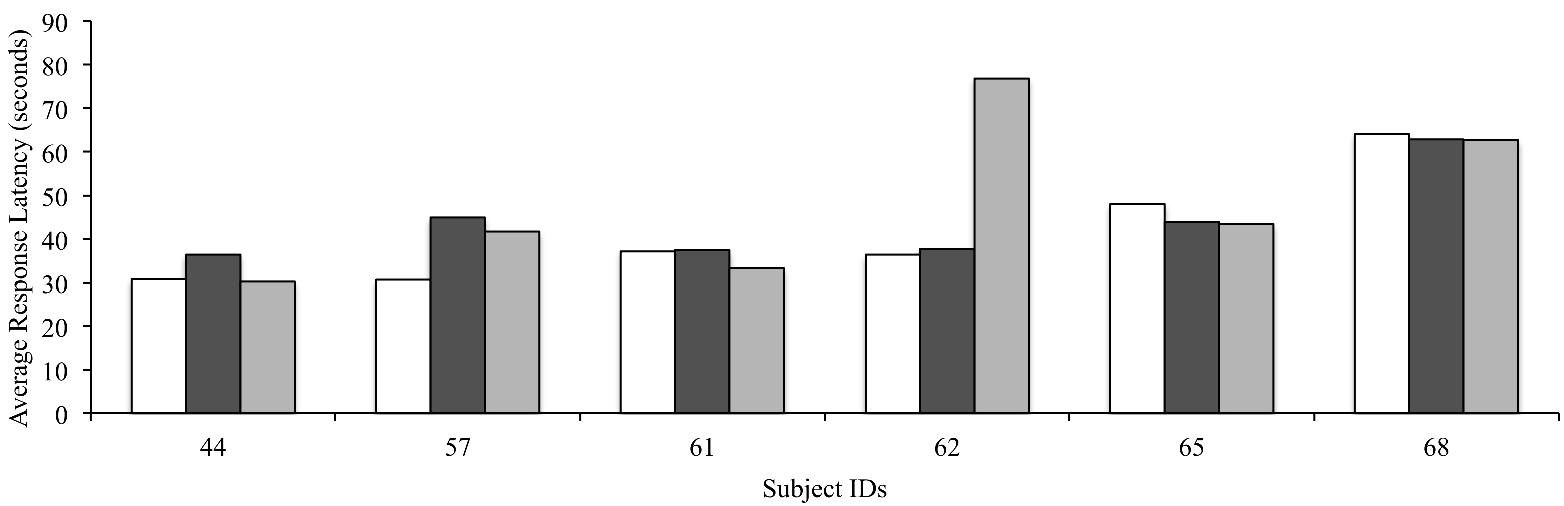


FIGURE 74b. Average response latency measures for all Subjects in the Endurance Group across Control, CD and CID conditions.

☐ Control ■ Contingent Delay ■ Contingent Interactive Delay

Experiment 3: Inattentive Group Subjects

Subject 45

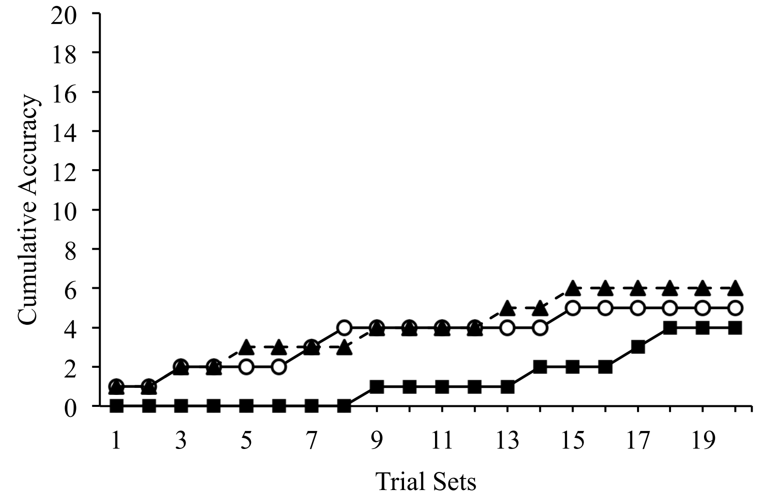


FIGURE 75a. Cumulative accuracy across Control, CD and CID conditions.

Subject 48

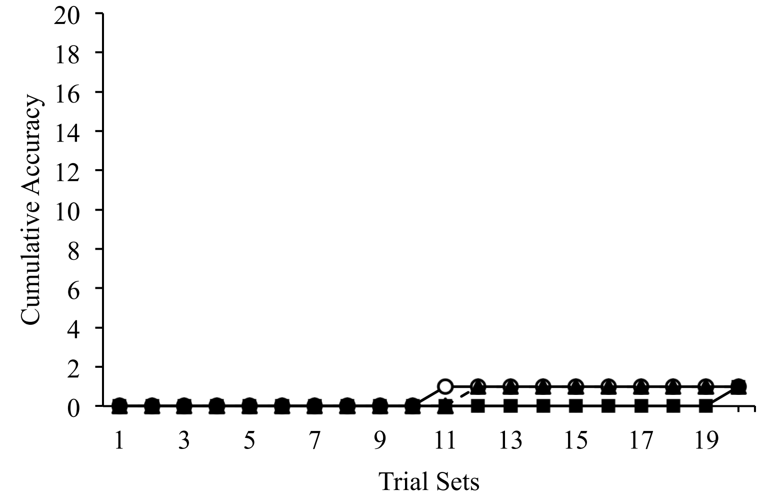


FIGURE 76a. Cumulative accuracy across Control, CD and CID conditions.

Subject 45

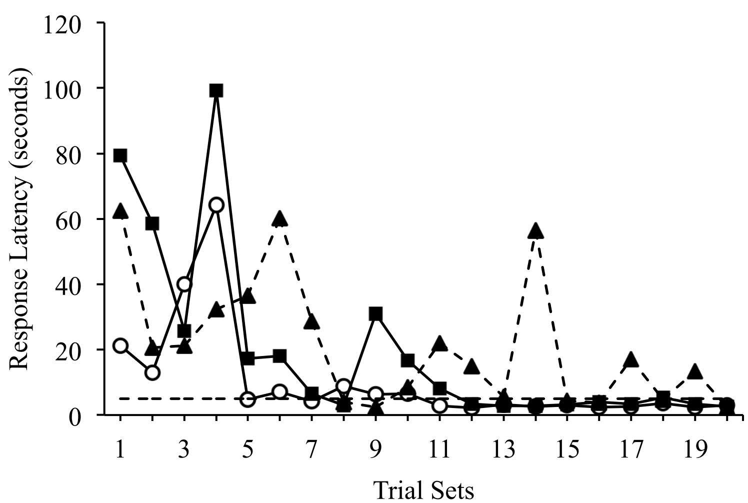


FIGURE 75b. Response latency measures across Control, CD and CID conditions.

Subject 48

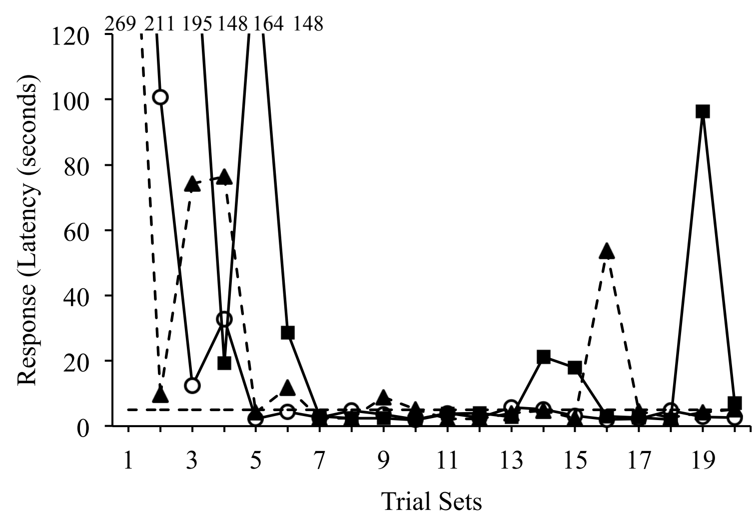


FIGURE 76b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 3: Inattentive Group Subjects

Subject 50

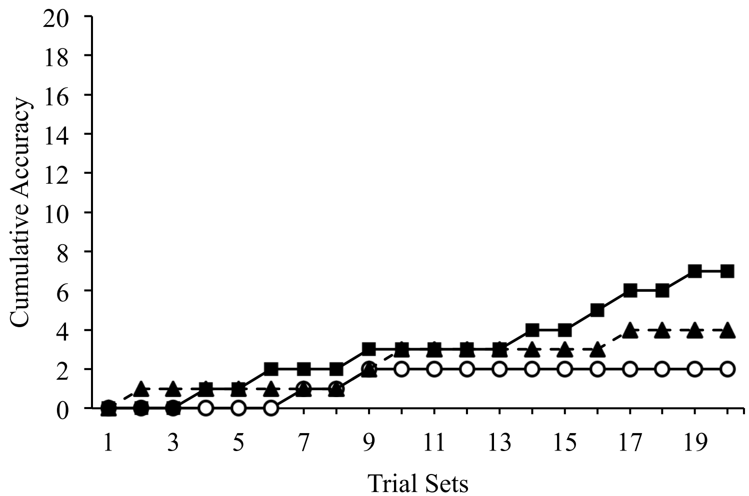


FIGURE 77a. Cumulative accuracy across Control, CD and CID conditions.

Subject 54

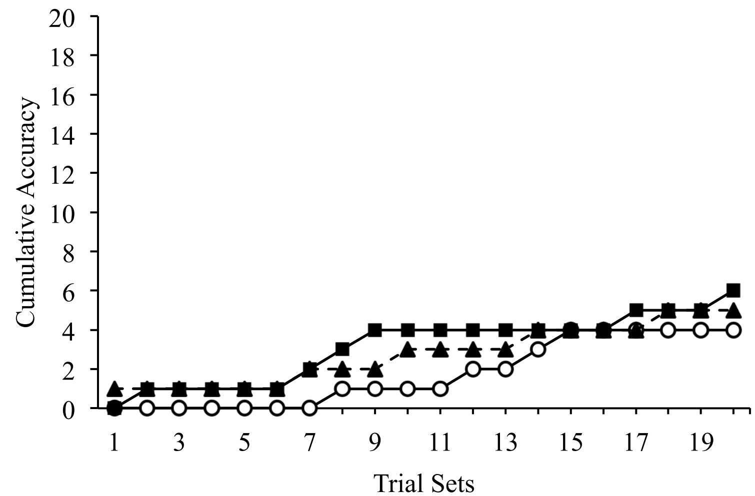


FIGURE 78a. Cumulative accuracy across Control, CD and CID conditions.

Subject 50

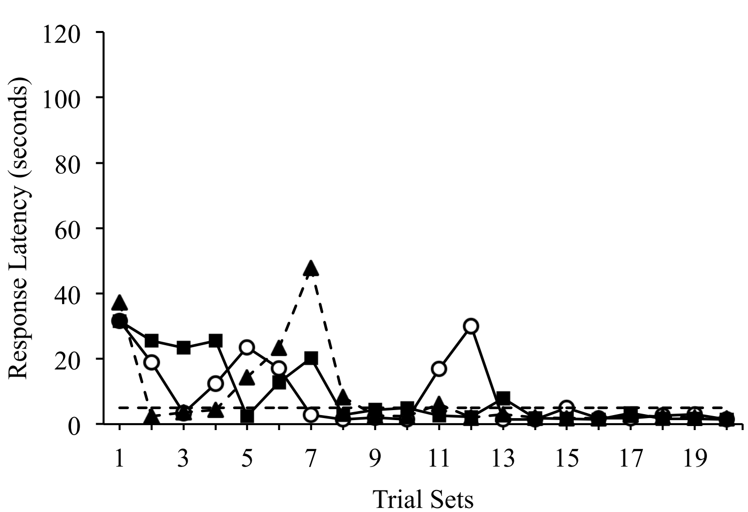


FIGURE 77b. Response latency measures across Control, CD and CID conditions.

Subject 54

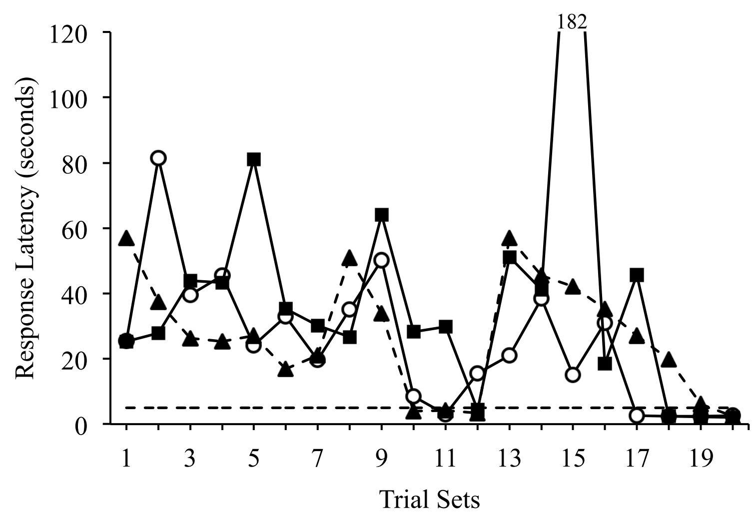


FIGURE 78b. Response latency measures across Control, CD and CID conditions.

Control — Contingent Delay — Contingent Interactive Delay — 5-seconds

Experiment 3: Inattentive Group Subjects

Subject 64

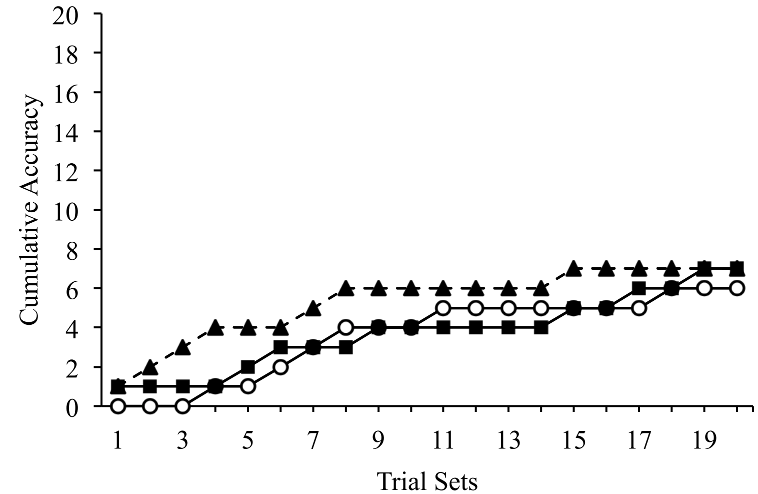


FIGURE 79a. Cumulative accuracy across Control, CD and CID conditions.

Subject 67

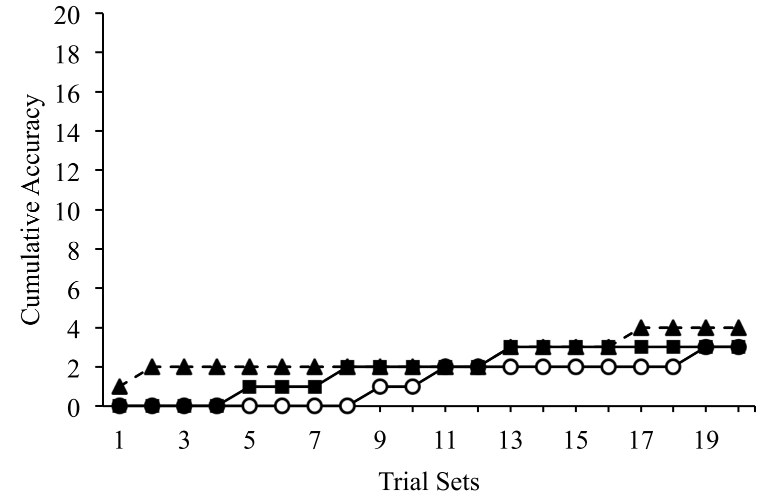


FIGURE 80a. Cumulative accuracy across Control, CD and CID conditions.

Subject 64

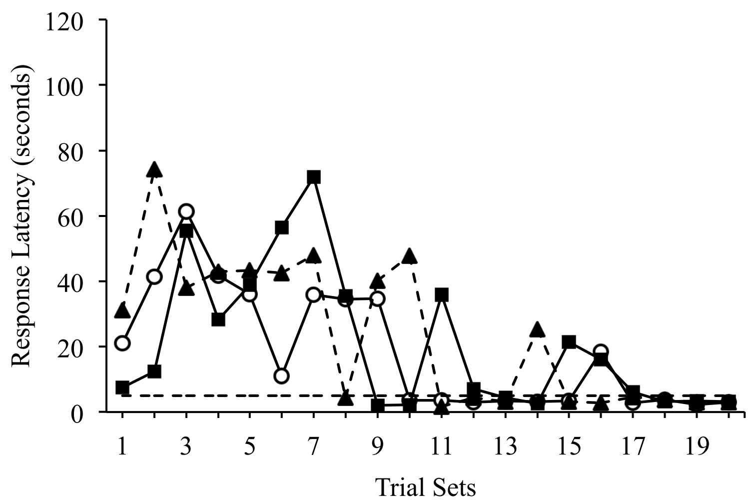


FIGURE 79b. Response latency measures across Control, CD and CID conditions.

Subject 67

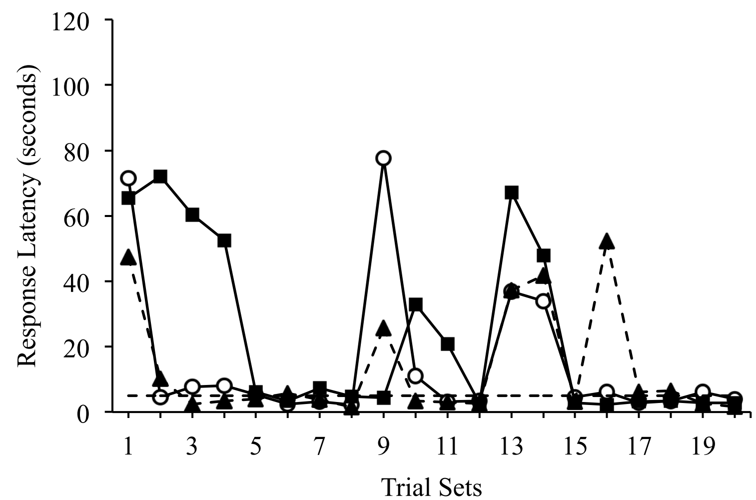


FIGURE 80b. Response latency measures across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay   5-seconds

Experiment 3: Inattentive Group Subjects

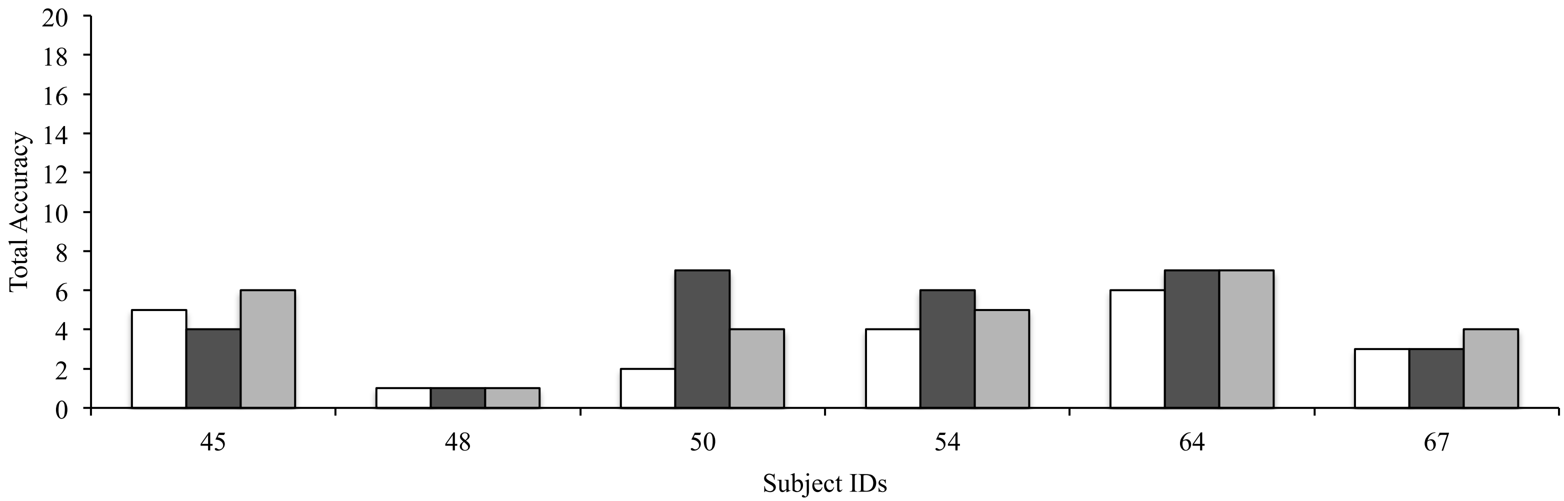


FIGURE 81a. Total accuracy of responding for all Subjects in the Inattentive Group across Control, CD and CID conditions.

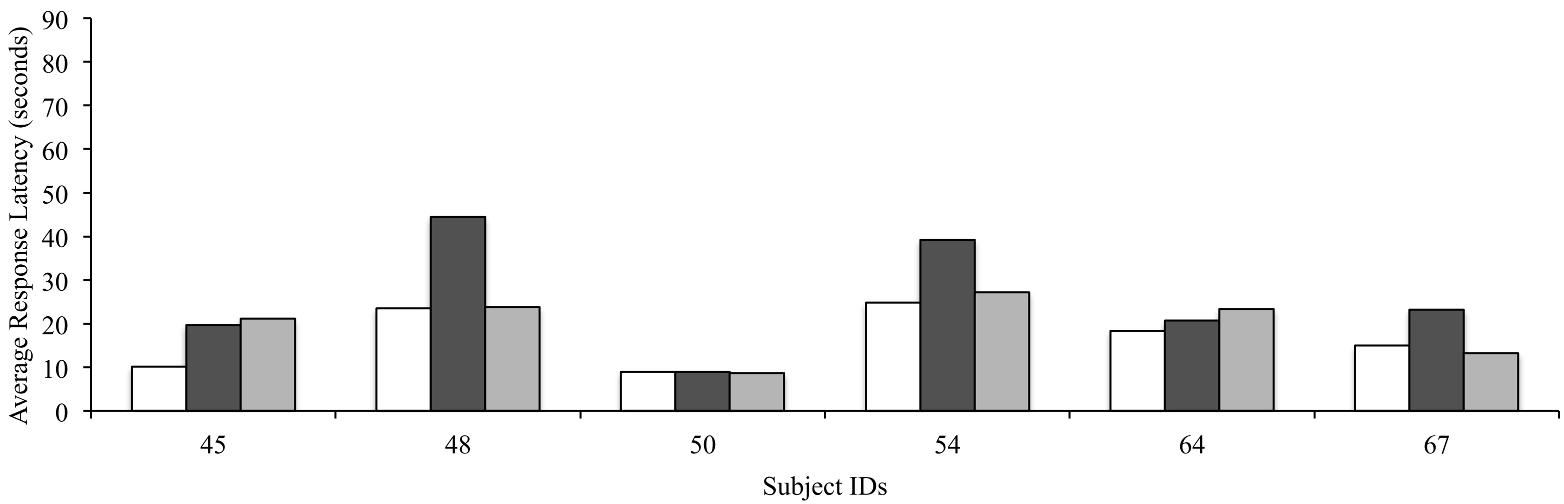


FIGURE 81b. Average response latency measures for all Subjects in the Inattentive Group across Control, CD and CID conditions.

Control   Contingent Delay   Contingent Interactive Delay

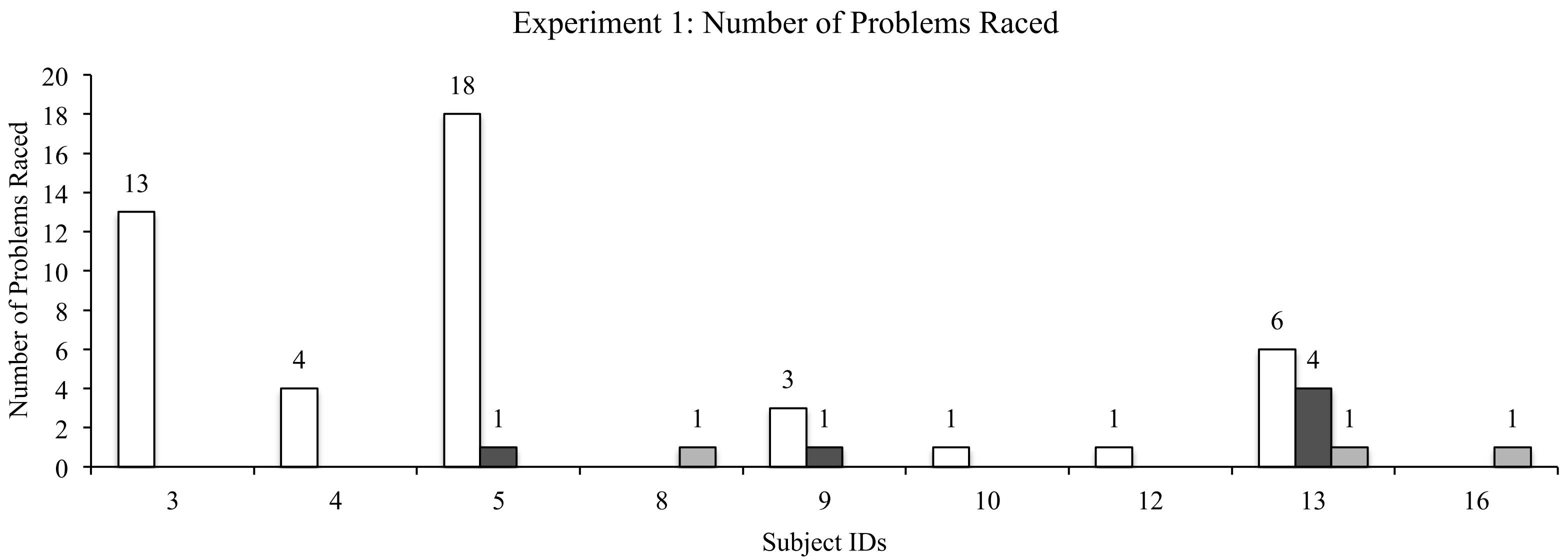


FIGURE 82a. Total number of problems raced for all Subjects with at least 1 response latency under 5-seconds across Control, CD and CID conditions in Experiment 1.

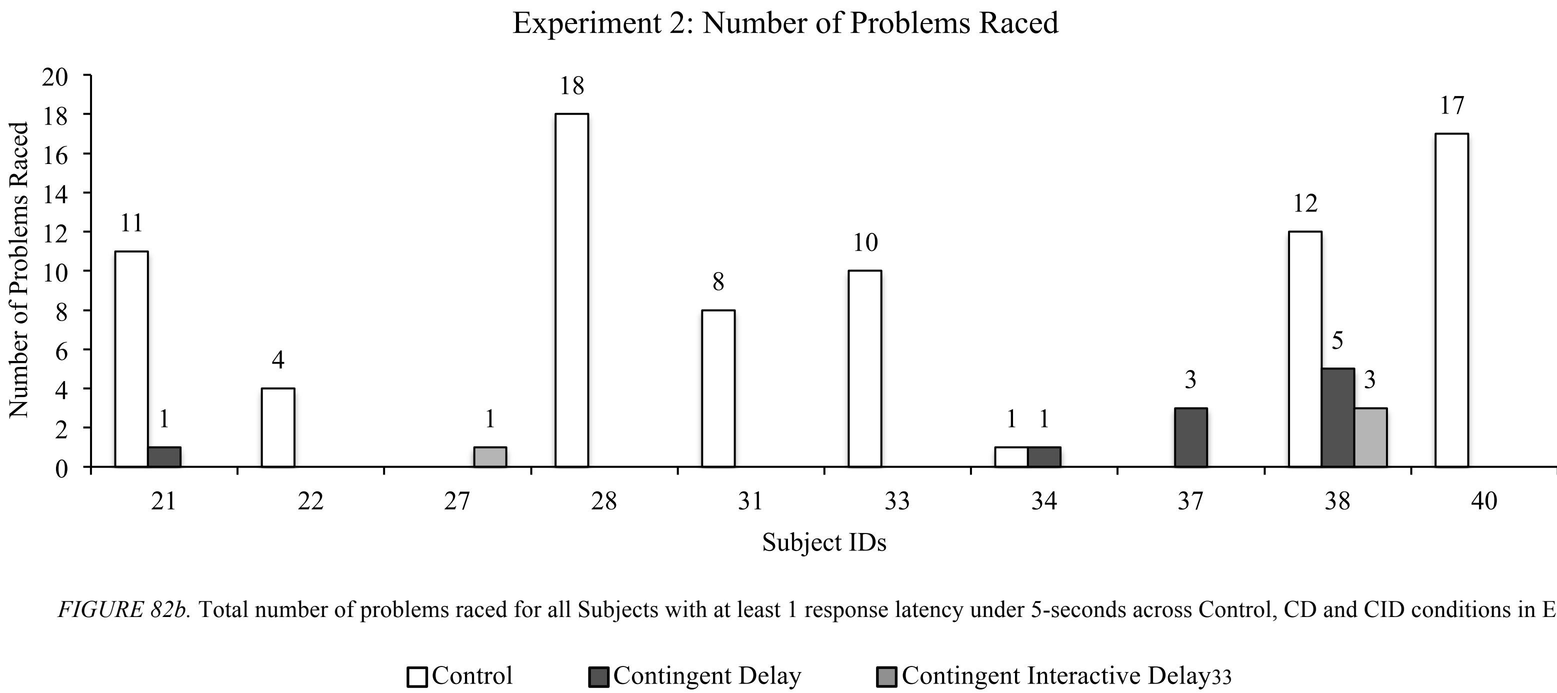


FIGURE 82b. Total number of problems raced for all Subjects with at least 1 response latency under 5-seconds across Control, CD and CID conditions in Experiment 2.

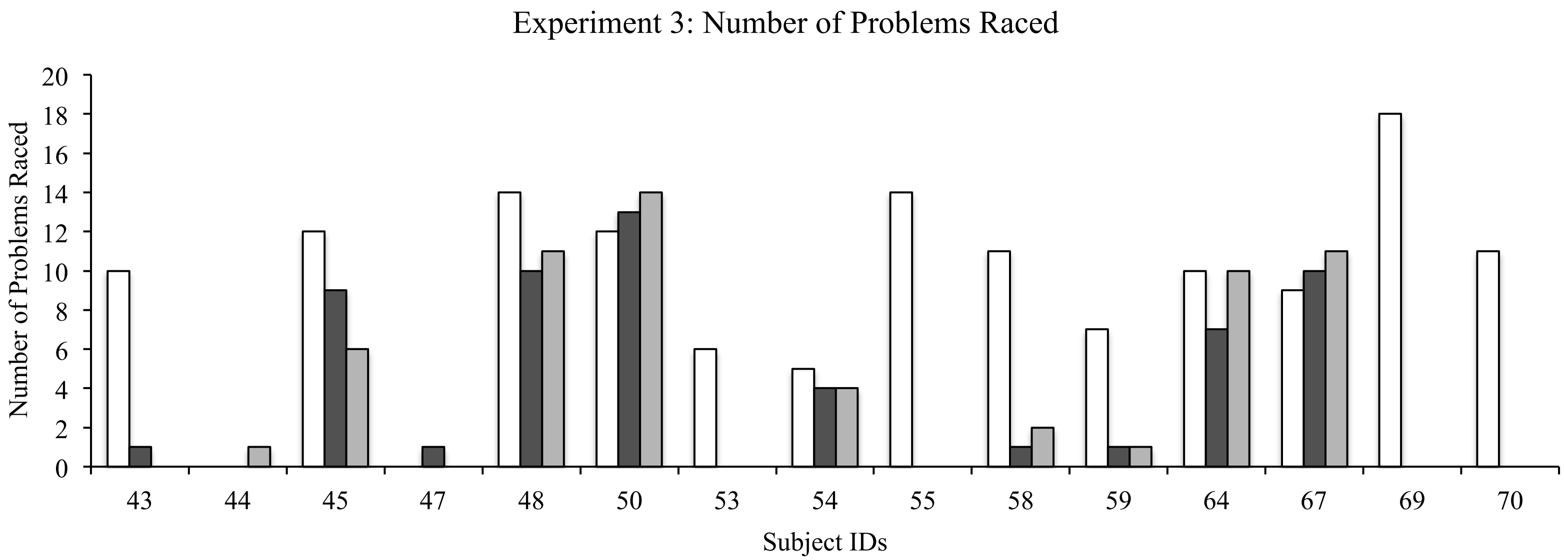


FIGURE 82c. Total number of problems raced for all Subjects with at least 1 response latency under 5-seconds across Control, CD and CID conditions in Experiment 3.

Control   Contingent Delay   Contingent Interactive Delay33

Experiment 3: Likert-Scale Ratings Given in Response to the Question: “On a scale of 1-9, how satisfied were you with the consequence for incorrect answers when the screen was [WHITE/BLUE/GRAY]?” (9 = “Very Satisfied” 1 = “Very Dissatisfied”)

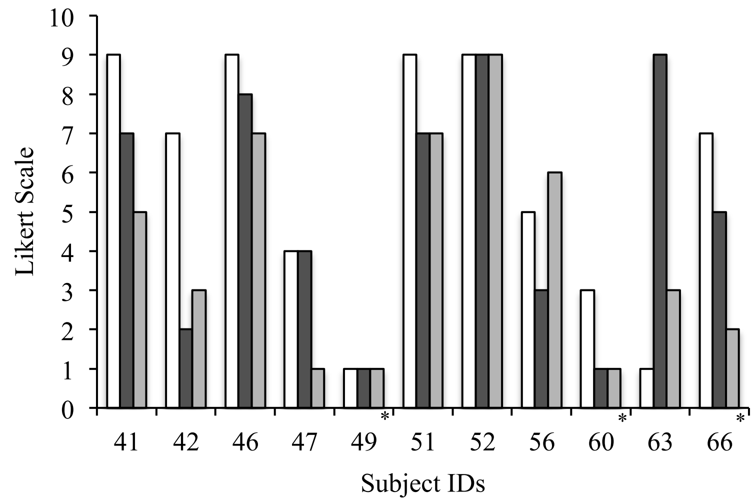


FIGURE 83a. Likert ratings provided by Proficiency Group subjects.

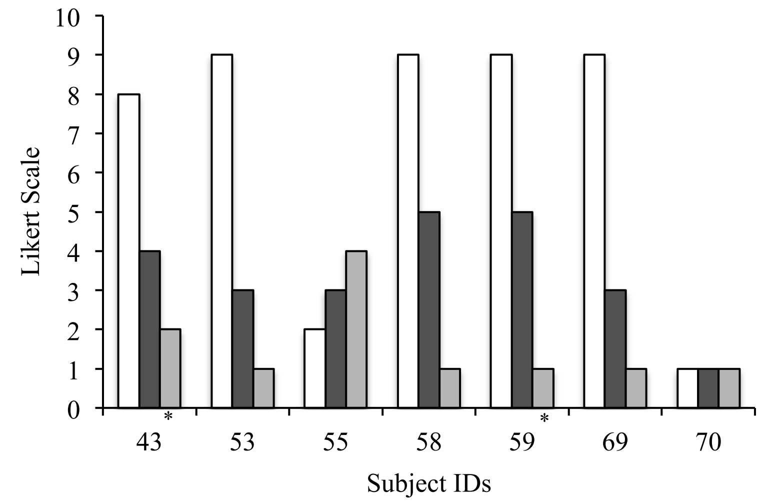


FIGURE 83b. Likert ratings provided by Avoidance Group subjects.

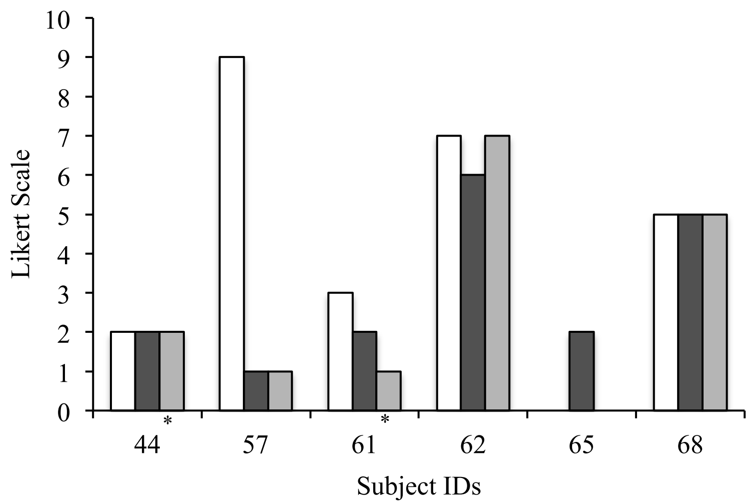


FIGURE 83c. Likert ratings provided by Endurancy Group subjects.

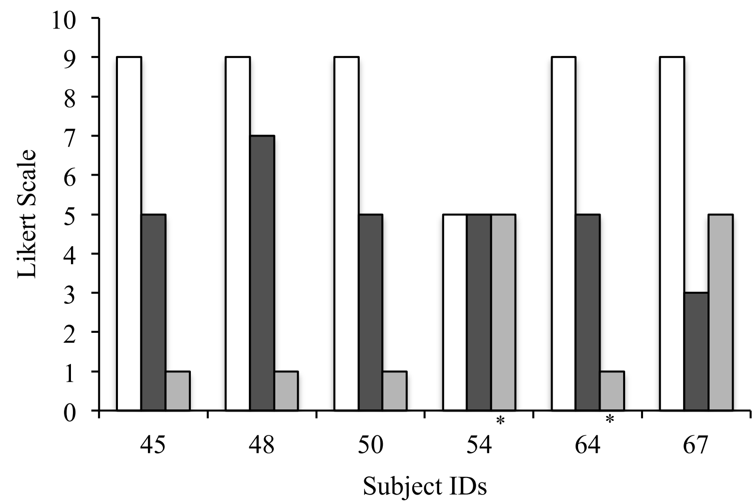


FIGURE 83d. Likert ratings provided by Inattentive Group subjects.

☐ White (Control) ■ Blue (Contingent Delay) ■ Gray (Contingent Interactive Delay33)

*Denotes the Subject choose an extra Contingent Interactive Delay condition problem to complete instead of an extra Contingent Delay problem during the forced choice condition.

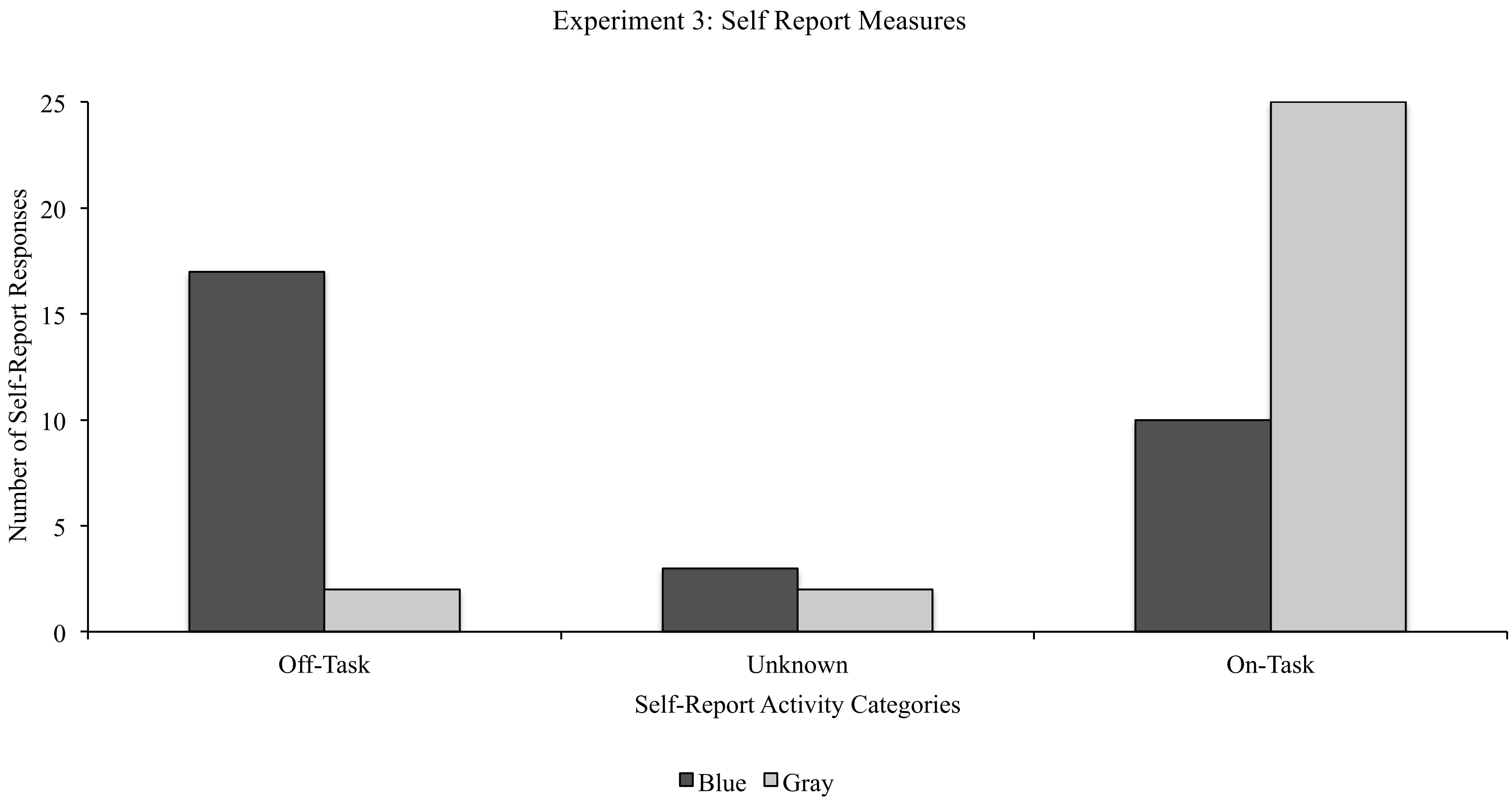


FIGURE 84. Categorization of self-report responses to the question, “If you answered a problem incorrect when the screen was [BLUE/GRAY] what did you do while you waited for the pause to end? Responses were categorized as “Off-Task” if subject indicated engaging in an activity not related to the computer task (e.g., “looked at my phone”, “watched T.V.”). Responses were categorized as “Unknown” if it was not clear what the subject reported doing (e.g., “nothing”). Finally, responses were categorized as “On-Task” if the subject reported engaging in activities related to the task (e.g., “watched for the button to change”, “checked my order of operations”, etc.)

Experiment 3: Self-Report Measures

Group	Subject ID	If you answered a problem incorrect when the screen was ____ what did you do while you waited for the pause to end? BLUE	GRAY
Proficiency Group	41	watched the numbers	watched the box under the numbers
	42	Checked my order of operations, corrected any mistakes and waited.	Checked my order of operations, corrected any mistakes and waited.
	46	Looked at my phone and kept checking to see if I could move on	Looked at the screen and waited
	47	I looked at my phone	Nothing
	49	click until it let me click it	click until it let me click it
	51	Tried to figure out what I did wrong on the problem.	Try to figure out what I did wrong while waiting for the button to appear.
	52	Thought how stupid this was and a waste of time.	Thought how stupid this was and a waste of time.
	56	Diverted my attention to something else	Watched the screen for the button
	60	Watched T.V	Constantly click where the button shows up and watch tv
	63	Looked over, or thought over the answers that I placed	Watched the screen for the move-on button
	66	I didn/'t get any wrong on blue. But you had to wait 30 seconds.	I waited until the button would change and say go on to the next problem.
Avoidance Group	43	Listened to music	Stared at it intently until the button changed.
	53	Texted on my phone.	I began to watch tv
	55	texted	watched for the button to change
	58	Jump rope	wait for the button
	59	I opened another a tab and browsed the internet.	I sat and tried to anticipate when the Click here button would appear.
	69	Stared at the screen	Kept clicking in the same spot
	70	Talked to friends on instant messenger, checked Facebook.	Watched nervously for the button to show up.
Endurance Group	44	Texting or changed the song I was listening to.	Watched the screen and waited for the cue to hit the button.
	57	Looked around and waited.	Watched for the click button to appear.
	61	sat there	watched for the button to say to push it
	62	Watched the screen	The first couple of time I thought it was going to go through the entire countdown so I missd it a few times. Then once I understood it was sometime during the 60 seconds I kept looking at the screen.
	65	Just stared at the computer, or looked around the room I was in.	I looked at the computer, to wait for the phrase to come up, so I can go to the next problem. Sometimes I would miss it, by staring off on accident.
	68	I counted down along with the computer screen.	I looked at the screen until the button would appear, so I wouldn't miss the button.
Inattentive Group	45	went on facebook	starred at the screen
	48	Watched a movie	Stared at the screen to make sure I did not miss the three seconds
	50	looked at my phone	stared at the button
	54	Nothing, I just sat and watched the numbers count down.	Just waited for the button to appear, I would pretend that the button would appear later and later and that the three seconds were to save the world. I was incredibly bored.
	64	Spaced out/watched tv until the timer went out.	I watched the screen and waited for the button to appear so I could move on to the next answer.
	67	checked e-mail in other tab, watched tv	watched extremely closely for button to change

TABLE 1: Self-report measures by group to the question, “If you answered a problem incorrect when the screen was [BLUE/GRAY] what did you do while you waited for the pause to end?”

Experiment 3: Self-Report Measures

Group	Participant ID	How did you solve the math problems in the study (e.g., in head, on paper, using a calculator, etc.)?
Proficiency Group	41	in head and sometimes with a calculator
	42	In my head for majority of problems, on paper for ones I answered incorrectly at first.
	46	In head and on paper when necessary
	47	I used a calculator
	49	on paper
	51	I solved them in my head.
	52	Using Wolfram Alpha
	56	About half and half between head and calculator. The first few on paper.
	60	calculator
	63	On paper and in head just depending on how the problem was set up
	66	In my head and also with a calculator.
Avoidance Group	43	A combination of in my head and using a calculator.
	53	In my head.
	55	in my head
	58	head
	59	In Head.
	69	In head
	70	In head.
Endurance Group	44	Head
	57	In head.
	61	on paper
	62	a mix between on paper and in my head
	65	in my head and on paper.
	68	On paper
Inattentive Group	45	in head
	48	on paper
	50	in my head
	54	In my head.
	64	In head and on paper
	67	in head/ on paper

TABLE 2: Self-report measures by group to the question, “How did you solve the math problems in the study (e.g., in head, on paper, using a calculator, etc.)?”

Experiment 3: Self-Report Measures

Group	Subject ID	Where did you complete this study?
Proficiency Group	41	library
	42	In the library.
	46	Bedroom
	47	Bedroom
	49	bedroom then living room watching tv
	51	My desk in my bedroom.
	52	Library
	56	Campus library
	60	bedroom
	63	Bedroom
	66	In my bedroom.
Avoidance Group	43	My bedroom
	53	On campus.
	55	bedroom
	58	dorm room
	59	Library on campus.
	69	Dining room
	70	Bedroom.
Endurance Group	44	Knowledge Center
	57	Library
	61	dorm room
	62	At work
	65	Bedroom
	68	The library
Inattentive Group	45	bedroom
	48	home
	50	in my room
	54	Dorm room.
	64	In my livingroom
	67	bedroom

TABLE 3: Self-report measures by group to the question, “Where did you complete this study?”