Precision teaching: Fast Practice or Merely More Practice Results in Better Learning?

A thesis submitted in fulfillment of the requirements for the degree

of

Master of Applied Psychology in Behaviour Analysis

at

The University of Waikato

by

Kong *Xiuyan* 孔秀颜

The University of Waikato

2009

Abstract

Doughty, Chase and O'Shields (2004) reviewed the Precision Teaching literature which they point out claimed that Precision Teaching is superior in producing learning outcomes and they posed concerns about the lack of well designed research to support these claims. One issue of concern was in much of the research Precision Teaching involved more practice than the teaching method to which it was compared. This present research aimed to compare the effects of fast practice and slow practice on learning three sets of 20 statistics definitions by tertiary students using a within-subject design. Following Doughty et al.'s (2004) suggestions, the amount of practice of the three sets of definitions was matched. All three sets were learned to some degree of accuracy before practice started. One set was practiced fast and another set was practices slowly and accurately over the same time period and until the fast practice set met a rate aim of 30 correct per min, the third set was practiced slowly and accurately for the same number of trials after this. There were non-timed tests of all sets of definitions prior to the start of any practice, after the rate aim was achieved, after the second slow practice finished and four weeks later. The amount of feedback for fast practice and the first slow and accurate practice was matched, but less feedback was provided for the second slow practice set. Three of the initial eight participants completed all experimental conditions. Results showed that response latencies to complete a response became shorted with both types of practice over practice but were shorter for the fast practiced items than for the slowly practiced items during the practice periods. The tests showed that accuracy was high after the extended practice and after four weeks no practice, regardless the method of practiced. The types and amount of feedback did not appear to have any effect on the learning outcome. Limitations that prevented a firm conclusion were discussed.

Acknowledgements

Upon the completion of this thesis, I am grateful to all people who have helped and supported me over this period. Thank you.

I'm lucky to have Professor Mary Foster as my supervisor who has always been patient, supportive and encouraging. I couldn't have come all this way without your support. Many thanks to Dr. Eric Messick who is always available when needed and who always asked about the project with interest. Thanks Mr. Andrew Malcolm for your patience and understanding in the communication for making the computer program. Thanks to other staffs in the department who have helped me in various ways.

I also want to thank all the participants for their time spent and effort made for this study.

Finally, I would like to thank my family in China for their trust and endless support. I'm also deeply thankful for Mr. Daniel Fung who had brightened my days in times of pressure.

爸爸，妈妈，家姐，祥华，涛姐，谢谢你们一直以来的支持和无条件的爱。

Contents

List of Tables

List of Figures

Appendices

Precision Teaching (PT), first introduced by Lindsley (1991), is a teaching procedure that has been designed to bring the learning outcomes of fluency (Doughty, Chase & O'Shields, 2004), which is considered to be the real mastery of a skill (Binder, Haughton & Bateman, 2002). The procedure in PT usually involves teachers setting a rate-aim for the students and the students make 1-min timed daily practices trying to be more fluent on the skill that is being learned in order to reach the rate-aim (Doughty et al., 2004). It also involves students plotting the number of correct responses they achieve per min on a standard celeration chart so that the teacher can easily monitor the students learning and make adjustments to teaching when needed. As Lindsley (1991) suggested, by looking at the celeration chart, the teacher can make immediate change to the teaching program if the chart shows no increase for some time. He also suggested that the celebration chart was also considered to provide a form of reinforcement because learners can easily see their improvement in rate of correct responses per min from the chart (Lindsley, 1991). Binder et al. (2002) suggested that skills learned under PT procedures sustain longer in our behaviour repertoire; and skills built upon those learned are retained longer and better. In other words, students would be able to remember skills better after PT and would be more ready to apply those skills to more complex skills. Doughty et al. (2004) summarised the learning outcomes of fluency as retention, endurance, stability, application and abduction (RESAA). In their definition of RESAA, retention refers to the ability to respond at the same rate after a period of no practice; endurance refers to the ability to sustain the high rate of responding for a longer duration (e.g., from 1 min to 5 min); stability refers to the ability to respond in the high rate in face of distractions; application refers to the ability to generalise and apply the skills learned under PT to other areas; abduction refers to the ability to perform an unlearned composite skill when its component skills have been learned with PT. Many who have used PT report that it results in faster learning, better memory for the materials learned, greater generalization and application

(Beneke, 1991; Bucklin, Dickenson & Brethower, 2000). These four outcomes are desirable in any educational setting. That is, to learn fast, to remember what has been taught at school, be able to generalize the learning to similar situations and, most importantly, be able to build upon that already learned to make new learning, as well as applying what was learned at school to daily life.

PT has been used in learning centres and remedial teaching programs for normal functioning individuals with learning difficulties. One organization that employs precision teaching techniques is The Morningside Academy in Seattle. The Morningside Academy combines precision teaching and direct instruction to teach under-achieving students who are labelled as "learning disabled" or are diagnosed with ADD or ADHD in general schools (Johnson & Layng, 1994). Johnson and Layng (1994) reported that Morningside Academy gave money-back guarantees to parents that all learners, including the so called "leaning disabled", "ADHD" and so forth, would reach beyond 65 percentile of the national achievement test before they transit into another school and so far Morningside had not given a refund in a seven year period up to the time of the report. Another example of using PT was the success reported from a reading program that incorporated precision teaching implemented by teaching assistants (Downer, 2007). Downer (2007) found that by adding 4-min daily practices over a period of time, students (year 1 to year 5) who previously had difficulty in reading increased in the number of words they were able to read in one minute, reaching the same level of the other peers who did not have the reading difficulties. When compared with other remedial programs, PT was shown to have similar gains immediately after the completion, however, the improvement gained from PT usually lasted longer (Le Grice, Mabin, & Graham, 1999; Singer-Dudek, & Greer, 2005).

Interventions that incorporated PT for individuals with some degree of disability have also been found to be effective, such as academic and physical skills training for children with

traumatic brain injury (Chapman, Ewing, & Mozzoni, 2005), mathematics and reading skills learning in children with neuropsychiatric diagnosis (Hartney, Mozzoni, & Fahoum, 2005), and basic functional motor skills training for adults with brain injury (Eastridge & Mozzoni, 2005). Thus there are data with children with learning difficulties and individuals with some level of disabilities supporting some of the claims of PT. Often, it is a gain in rate that is studied but not the other fluency "outcomes" (RESAA).

There are also a few studies of the use of PT with tertiary students. For example, Beneke (1991) who used PT to teach reading skills to introductory psychology students and found that by practicing to read a passage and recall it quickly for less than 5 min daily over a semester, students gained a 45% increase in reading rate and a 75% increase in recall rate for the readings; whereas a control group who did not participate in the study showed no increase. In another study with college students, Bucklin, Dickenson and Brethower (2000) taught arbitrary relations between Hebrew symbols, nonsense words and numbers with PT and found that, when compared to training to accuracy only, relations directly learned under PT were retained longer even after 16 weeks. Moreover, relations not directly learned in PT but built upon those learned were responded to more accurately and retained longer without practice. In spite of showing that PT had the potential of improving learning, there is not much research on its outcomes.

"Practice makes perfect", as the old saying suggested, practice alone can help mastering a skill. Driskell, Willis and Copper (1991) pointed out that overlearning, defined as extensive practice after accuracy, could produce better long term retention for cognitive tasks. That is to say, practice alone has a positive effect on retention. Doughty, Chase and O'Shields (2004) reviewed the literature on precision teaching and found that many of the studies claiming precision teaching to be superior to other methods of teaching compared one group receiving precision teaching with another group received normal schooling where there was

far less practice required from the students. Examples are the Morningside Academy (Johson and Layng, 2004), Bucklin, Dickinson and Brethower's (2000) study using equivalence classes and Beneke's study on reading skills (1991). In these examples, students/participants in the rate-building group had more practices than their peers who did not use PT. Therefore, it was unclear whether the learning outcomes reported from those studies resulted from the fast and extensive practices or merely from the extensive practices. The small number of studies that have tried to control the amount of practice and reinforcement did not find the acclaimed superior effects. For example, Darvel and Rose (2006) had adults with developmental disabilities practiced reading quickly or slowly and accurately the same number of times and found that there was no differences in the outcomes of percentage correct for acquisition, maintenance and generalization between the two method of practices. Peladeau, Forget and Gagne (2003) implemented a computerized flashcard teaching program in a college quantitative research method class to give participants opportunity to practice materials the same number of times either under a rate-building condition or an accuracy-only condition. A between-subject comparison showed that the two groups of participants who had extensive practices after accuracy had better grades in the course as well as having better retention score than the group who did not have the extra practice. The two groups who had the same amount of practice, however, did not differ in their course grades or long term retention score. Fox and Ghezzi (2003) also ran a between-subject research study using a computer program to examine the effect of rate-building and non rate-building practices on learning outcomes. The authors reported that the low power of the study resulted from using between-subject design prevented a conclusion on the differences between fast and slow practice; and data collected by the computer program used in their study did not allow for examination of whether the students responded quickly to the sets of items for slow practice. The methodology of such studies normally requires fast timed-practices for the rate-building

practice; the other practise may involve participants being told to be as accurate as they can be, and they may also be told not to worry about speed. The problem is that although the methodology aims at separating speed from amount of practices, participants could still respond fast to each item in the slow or paced practice condition resulting in a lack of distinction between the types of practice. This prevents firm conclusions from these studies. Nonetheless, they suggest a means of systematically presenting stimuli and measuring response latency in research in the precision teaching area.

Doughty et al.'s (2004) review of PT drew attention to the question of whether both speed and repeated practice are necessary or whether the mere extensive practice alone is sufficient to produce the benefits claimed for PT in the precision teaching literature. According to their review, the main concern arising from the research was the lack of control of the amount of practices and reinforcements in precision teaching and in the comparison teaching method. They suggested there are three components a research study should include in order to separate the effect of speed and practice on learning outcomes. First, the investigation should use a within-subject design. This an issue for Peladeau et al.'s (2003) and Fox and Ghezzi's (2003) study. Peladeau et al. suggested that their students differed in ability and so it was difficult to make a between-subject comparison since learners would be at different starting points hence a within-subject design would have been preferable. Second, the investigation should include two different skills or subjects for learning, one for rate-building practice and the other for non rate-building practice to prevent practicing one skill or one subject with both methods. Thus the material needs to be divided in some way to try and prevent practice of one skill or set of facts from facilitating the learning of the other skill or set of facts. Third, the investigation should also equalise the amount of practices and the amount of reinforcement for both types of practice. Doughty et al. (2004) suggested that the feedbacks given to participants at the end of practice could be treated as reinforcers. That is,

they suggested that participants' behaviour was maintained by the feedback on the number of correct responses, or the rate of correct responses and/or the graphs of the data. Since this review, three unpublished theses (Wheetley, 2005; McGregor, 2006, & Clark, 2007) have attempted to solve some or all of these issues and they will be discussed in detail below.

Wheetley (2005) examined the effects of extended practice on a matching-to-sample task under both rate-building and non-rate-building conditions on retention, endurance, stability and application. This study employed a within-subject and reversal-treatment design. Participants learned to match 3 sets of English letters to Japanese characters using touch screen computers thus each response and its response latency was automatically recorded by the computer. Each participant went through two methods of practising: one being fast and accurate, the second being accurate only and each experienced one of these twice. Each practice method used one of the 3 English letter and Japanese symbol sets. In the experiment, participant were shown an English letter in the middle of the computer screen and the 7 Japanese symbols in that set arranged in a circle around the letter and they were required to touch the symbol that corresponded to that letter. There was an error correction procedure in which the correct response was first bolded and set in red font to prompt correct responses before an unprompted response was required. Sessions for each condition began with an error correction procedure to show the 7 paring stimuli in each set then followed by a 20-s practice drill in the accuracy and rate condition and a 7-response practice drill in the accuracy only condition; the error correction procedures were interspersed with the practice drills. Feedback on whether a response was correct or incorrect was provided by the computer for each response in the error correction procedure only; no such feedback was given for the responses in the practice drill. At the end of each 20 s practice drill of accuracy and rate-building practice, the computer showed the number of correct and incorrect responses; and such information was provided after every 7- response practice drill for the accuracy-only set. The

number of presentations was made the same between the accuracy and rate-building set and the accuracy-only set. One-second delays were imposed between the presentation of the English letter and the 7 Japanese symbols for the accuracy-only sets to ensure that the response rates differed across the practice sets. As reported in the results, there was no differentiation in the rate of responding (average response latencies) between the accuracy and rate-building and the accuracy only condition at the end of the training (last 40 responses) for 3 out of 4 participants. Particularly, 1 participant who started with accuracy only condition without experiencing any rate-building practice shown low average response latencies for the last 40 responses, similar to that of the accuracy and rate-building condition. This suggests that the different instructions for the two conditions did not result in the intended difference in rate of response. Responses latency was not examined in the test in her study. The test results for retention, endurance, stability and application were mixed: accuracy and rate-building were better for some participants in retention while accuracy only was better for the others. Limitations suggested by Wheetley (2005) were that a longer interval was needed before the test of retention and also that an improved fluency criterion was needed. There was one further limitation, that is, the study did not control for the amount of feedback provided for each practice set. For the accuracy and rate-building set, participants were shown their results at the end of every 20-s practice drill whereas for the accuracy-only set they were shown the results after every 7-response practice drill. Doughty et al. (2004) suggested that such feedback could have served as reinforcers. Thus, if so, there could be different amounts of reinforcement between these two practice methods. Different rates of reinforcement had been found to produce different learning outcomes (Walker, 2008), therefore, this further prevents a firm conclusion from the study.

McGregor (2006) also tried to separate the effect of rate and accuracy on retention. She taught 7 participants 10 multiplication facts and then had them practice half the facts in a

rate-building manner and the other half in a rate-controlled manner. Practices in the rate-building manner involved 1 min fast practice with a rate aim of 70-100 correct responses and less than 2 errors per minute; participants controlled the rate of presentation. In the rate-controlled practice method, there was a 3 s delay between presentations so that participant could not build up their response rates as they could with the rate-building method; the rate of presentation was controlled by the computer. In the study, all facts were learned to an accuracy criterion and practiced slowly the same number of times for each participant before the rate-building practice in which the participant controlled the rate of presentations (termed by Lindsley (1992) as free-operant productions) and the rate-controlled practice in which the computer controlled the maximum rate of presentations. Participants had up to three daily sessions of 1 min rate-building practice followed by a rate-controlled practice yoking the number of practices in the previous period of rate-building practices. The amount of feedback was also equalized between the two methods of practice. Participants were tested in 1-min timing for each set of multiplication facts although they were not required to do the tests quickly.

Results of McGregor's (2006) study showed that accuracy was high at the end of both practices and remained so in the 4 and 8 weeks post test; practicing the facts fast or slowly did not have a different effect on accuracy up to 8 weeks after no practice. However, all participants reached the accuracy criterion faster in the rate-controlled practice than in the rate-building practice. Furthermore, the rate of responses differed between the two sets of facts at the end of practice, but this difference decreased over time (at 8 weeks). Although the rate of correct responses in the tests were generally higher for the facts practiced in the rate-building condition than those practiced in the rate-controlled condition after 4 weeks, this difference decreased to a minimum after 8 weeks. The greatest decrease in rates was for the facts practiced in the rate-building condition as they were faster to start with. McGregor

(2006) concluded that responses rates for both sets of the facts decreased over time without further practice, regardless whether the facts were initially practiced in a rate-building or rate-controlled manner.

One issue with McGregor's study was that the tests were timed for 1 min, which was the same as the time used in rate-building practices. It is unclear whether this had influenced the rate of responding in the assessment. More specifically, that tests were timed may have prompted participants to response quickly. A close look at their rate of response data revealed that the rate of responses for 1-min timed tests with the 2 times table, which was not extensively practiced in any sessions was high at the beginning of the practice phase and, interestingly, remained high through all post tests. This suggests that having timed 1-min tests did encourage participants to respond fast and gave them practice at responding fast with the different sets of definitions. An un-timed assessment method is needed to help clarify this. Further, although it was shown that the rate of response was slower in the rate-controlled practice with the 3 s enforced delay between presentations, this did not show whether participants responded quickly or slowly to each individual item (i.e., whether they had short or long response latencies to complete the task); and this could only be answered by examining the latencies for the responses to each item.

In the last thesis, Clark (2007) taught 4 school children 80 times table facts using flashcards. She used a within-subject design. Participants were first taught 80 unknown times table facts slowly with immediate correction to inaccurate responses. The 80 facts were then divided into 2 sets of 40, one set for rate-building practice and the other for non rate-building practice that focused only on accuracy. Flashcards were presented by the experimenter in the practice sessions and in all tests. In each daily session, participants first completed two 1-min fast (rate-building) practices on one set of facts followed by 2 slow practices on the other set of facts yoking the total number of trials in the previous two 1-min fast practices. The

maximum duration between the presentation of a question and a response for fast practice items was 3 s whereas it was 6 s for slow practice items. In both fast and slow practice, participants were provided with feedback on whether the responses were correct or incorrect by the experimenter. There was an error correction procedure, which wwas the same as the teaching procedure, after each practice to ensure that every item had been correctly responded to the same number of times. Participants were told the rate per minute at the end of each fast practice and they were also told their accuracy score at the end of each slow practice. After the four practices, half of the 80 items were tested in a way that was the same as the fast practice; it was followed by a test that was the same as the slow practice on the other half. For the daily tests, all 80 facts were mixed together and shuffled before dividing into two sets; and this procedure was done on every practice day. Unlike McGregor (2006) and Wheetley (2005), Clark (2007) employed an accuracy aim of 95% correct over 3 consecutive days or 6 weeks after experiment started instead of setting a rate aim as the completion criterion. Participants were then tested 1 day, 2, 4, 8 and 12 weeks after practice had completed on maintenance (retention), stability, generalisation and application.

Clark (2007) found that participants responded at a much slower rate for the fast practice than they did for the slow practice during acquisition. However, the difference decreased over time in follow up tests after practice. Accuracy was generally higher for the slow practice items than for the fast practice items throughout the practice period but it was similar between the two methods in all follow-up tests. Accuracy did not reach 100% correct for all participants during practice but increased in the follow-up tests in which it reached 100% correct. The author also reported that no difference between fast and slow practice facts was found in follow-up tests for stability, generalization and application. Clark (2007) concluded that when the amount of practice and feedback was held constant, fast (rate-building) practice had advantage over slow (non rate-building) practice only in the rate of response but not in

accuracy during the practice period; rate-building practice produced no better accuracy or faster ate of response in follow-up tests in maintenance (retention), stability, generalisation and application.

Clark (2007) commented that several factors prevented a firm conclusion in her study; these were the learning of times table facts in class, the school holiday occurring during the research and the possibility of bias resulted from the method of testing. However, there was one serious problem in the testing procedure that may also have affected the results. In the daily testing procedure, facts from both rate-building and non rate-building practice sets were mixed and then divided into two equal sets which were tested under testing procedure that were similar to either the fluency training condition or the accuracy training condition. As a result, the slowly practiced facts may have been practiced in a rate-building manner in the daily tests which means that the intended difference between the two methods may have been lost. Although the results in Clark's (2007) thesis showed that the rate of correct responses for the slow practice set during the practice period was almost half of that of the fast practice set, the rate of correct responses was calculated by dividing the number of correct responses by the duration. This rate of correct response reflected more about the rate of presentation of the items rather than the actual latencies taken to complete each response. In other words, it is not clear whether the overall latency for completing a response was different enough between the rate-building and the non rate-building practices to produce possible different comes. One last problem is that the experimenter providing feedback for each response, especially for the fast practice set, may have prevented participants from responding quickly enough, up to the level that is required to produce the retention of high accuracy.

All three theses above used a within-subject design and equalized amount of practices to compare the outcomes of rate-building practice and of rate-controlled practice. All had items learned to some degree of accuracy before differential practice started. Although not

stated directly in their theses, this learning to accuracy procedure may be included to prevent participants practising the incorrect responses over and over again (Howell & Lorson, 1995). Two of them, Wheetley (2005) and McGregor (2006) employed a rate aim while one, Clark (2007), employed an accuracy aim. Precision teaching typically trains students to be both accurate and fast and requires students to reach a fast rate of responding in order to reach fluency (Binder, Haughton, & Bateman, 2002). Students in a study that employs an accuracy aim for completion criterion may not reach the rate requirement for optimal outcomes. For example, in Clark's (2007) study, none of the participants reached 100% accuracy when they met the completion criterion; and the rate of responding for the fast practice set was much lower than the published 70-100 correct per minute for times table (Binder et al., 2002). Binder et al. (2002) pointed out that a rate aim that has been established for certain skill may be desirable for obtaining the optimal outcomes from rate building practice. But for the skills that there is no published rate aims, a hard but obtainable rate aim is desirable (Pocock, 200?). The three studies also differed in that two of them used computer to present the stimuli sets for both fast and slow practice whereas one study used experimenter manually presenting the stimuli. Computer presentations have the advantages of giving consistent presentations for all stimuli in different conditions and being accurate in controlling times for presentation. It could also help in recording the actual response latencies for each item. All three studies used within-subject design as recommended by Doughty et al. (2005). However, exposing the same participant to both rate-building and non rate-building practise, usually with a short delay between two types of practices and sometimes almost at the same time, may result in less differential responding to the two ways of practise. More specifically, participating in rate-building practice may increase the speed of responding in non rate-building practice, i.e., there may be some carryover effect of the practise between sets of facts or skills as pointed out by Pocock (2006). Up to the time when this study was conducted, no studies had

examined whether there would be carryover effect from rate-building practice to non rate-building practice when comparing these two methods using within-subject design. If participants had responded faster in non rate-building practice as a result of being exposed to the rate-building practice, then the similar outcomes found from these two methods of practice in the three theses mentioned above may be due to the lack of difference in the response latencies between these two methods. Finally, as pointed out by Clark (2007), the testing procedures employed in the above studies were advantageous for one or the other practice method and efforts must be made to minimise the influence of testing on the results of practice. Since in a within-subject design, participants will go through both styles of practices, the bias in the testing procedure appears to be difficult to eliminate: testing in a timed manner is not only biased towards items practiced under the rate-building method but may also increase the speed of responding to the slowly practiced items by providing fast practice in the tests. Thus, testing in a slow (un-timed) manner should help prevent participants practicing responding quickly to the slowly practiced items. Testing in an un-timed manner is probably preferable so it should not affect the slowly practiced items and nor does it prevent fast responses to the items practiced quickly.

To help clarify the effect of different types of practise requires research in which the comparison practice methods include different rates of responding but in which the amounts of practice and feedback are equalized; it should include a published or difficult but achievable rate aim for the particular skills being practiced for optimal learning outcomes; and it should employ an appropriate testing procedure. Examination of the differences in response latency in the comparison practice methods is also needed for two reasons. One is that it can tell whether there is clear distinction in the latencies for completing a response between the two practice methods. Second, it can help determining whether the response latencies in the rate-control practise are influenced by those in the rate-building practise when

these two are paired in a within-subject design. A learning to accuracy procedure before the differential training should also be included to prevent participants from practicing errors over and over again (Howell & Lorson, 1995).

To achieve this, current study employed a within-subject design and had each participant practice in fast timed practices and in slow rate-controlled practices with different sets of statistics definitions. Participants learned all definitions to a certain level of accuracy before they started the training. Once accuracy was obtained, the definitions were divided into 3 sets. A rate aim of 30 correct responses per minute was used as suggested by Merbitz, Vieitez, Merbitz and Binder (2004). The training was arranged so that in the first part (Practice I Phase), each session consisted of 3 fast and timed 1-min practices on one set of definitions (fast practice; FP) followed by 3 blocks of slow and rate-controlled practices (slow practice 1; SP1) on another set of definitions. The amounts of practice and feedback were equalized in this part of the training. After participants reached the rate aim in the first part of the training in FP, they moved onto the second part of training (Pracitice II Phase) which included only slow and rate-controlled practices (slow practice; SP2) on the third set of definitions. The rate of presentations of items in SP2 was identical to SP1. SP2 was conducted without any fast practices happening at the same time, so it was hoped that it would not be influenced by FP even if SP1 had been. Adding SP2 in the procedure allowed for examination of how accuracy and response latency would change at the end of Practice I Phase for the set that had not been practiced; and it allowed for the assessment of the latencies of slow practice without the influence of fast practice. If response latency decreased in SP1 as well as SP2, FP may not be the sole influence for the decrease seen in SP1. Participants were tested for accuracy in an un-timed testing procedure but response latencies were monitored during acquisition and in all tests.

With the above experimental arrangements, it was hoped to be able to discover whether there were differences in accuracy and response latency between these two methods of practice during the practice period (Practice I & II), immediately after practice and up to 4 weeks after practice. Based on Wheetley (2005), McGregor (2006) and Clark (2007)'s findings, it was thought that there would be no differences in accuracy between the practice methods. The design also allowed a more detail examination of the response latencies to see whether the lack of difference in the results in accuracy, if found, were due to the lack of difference in response latencies between the two comparison methods of practice.

Method

Participants

Recruitment advertisements were posted on the psychology department notice board and handed out at the end of a first year psychology class and at the end of one pre-university preparation class in psychology. Eight students (one male and 7 females) responded to the advertisement and indicated they wanted to participate after meetings with the experimenter for briefings. Five of the participants were from first year psychology courses, 2 of them from the second year and 1 from the pre university preparation course. English was the second language for one of the female participants but it was the first language for the remainders. Figure 1 shows the number of participants in each phase of the study; only 3 of the participants completed all phases of the experiment. Participants from the first year psychology course received 4 points towards their course credit for that course.

Materials

Sixty definitions of statistical terms were selected from the statistics text book *Fundamental Statistics for the Behavioral Sciences* by David C. Howell (2008). The

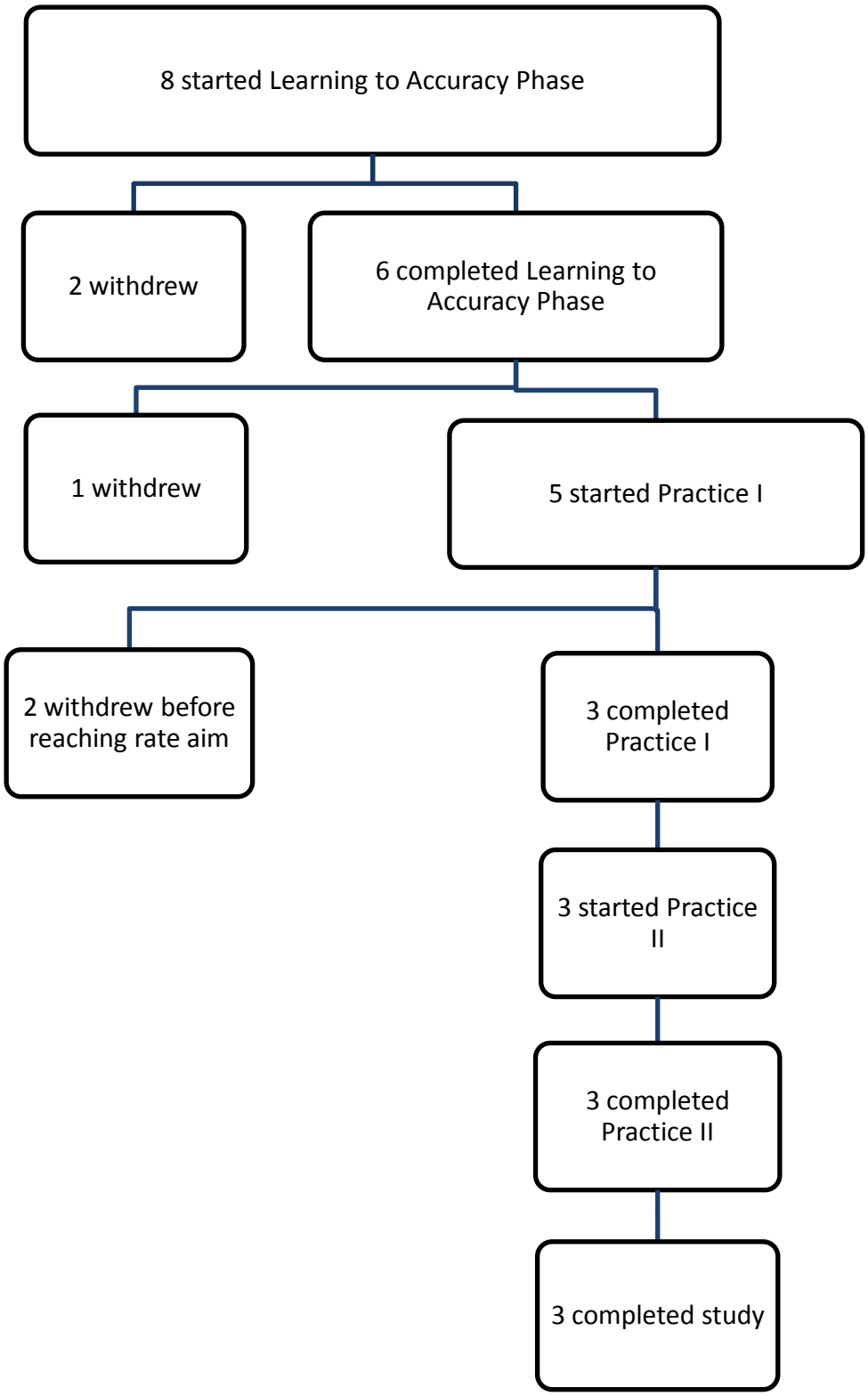


Figure 1. Chart showing the number of participants at each phase of the study.

terms were selected because they would be useful for students taking the psychology courses. Key elements for each definition included in the learning were pinpointed with consultation with a statistics teacher.

Sixty laminated paper flashcards (3.5cm×6cm) with a statistics term (font size 20) printed on one side and its definition (font size 14) on the other were used in the Learning to Accuracy Phase of the experiment. Appendix 3 shows the actual sizes and contents displayed on the cards.

A desktop computer with a specially designed program was used in the practice phase of the experiment. The monitor was used to present terms and definitions, and the keyboard, which was used for participants to indicate they were responding.

Early sessions of the Learning to Accuracy Phase took place in a room (size: 2.5m× 2.5m) in the psychology department with a large table and two chairs. However, this room later became unavailable and the remaining sessions of the Learning to Accuracy Phase and all the sessions of the Practice Phases took place in another room (size: 3m×2.1m) with two desks, two chairs and one computer with monitor, keyboard, mouse and CPU on one of the desk. The door was shut when a session started and all sessions were one to one sessions with only the participant and the experimenter in the room.

A video camera was used in the practice phases (Practice I and Practice II) of the experiment to record presentation on the computer screen and the responses from participants for inter-observer reliability checks.

Semi-log charts (Appendix 8) were used for participants to graph their performance in fast practice (FP) and a chart for accuracy (Appendix 9) was used for slow practice 1(SP1).

Procedure

After the participants had responded to the advertisement and been given information about the study, an initial meeting time was arranged.

Information sheets (Appendix 1) about the study and how much work and time it would require were handed out and were read out loud by the experimenter to each of the participants in their first session. Participants were then given time to read through the information sheet and they were also given the opportunity to ask questions. Consent forms (Appendix 2) were collected.

The procedure consisted of three phases: Learning to Accuracy Phase, Practice I Phase and Practice II Phase. These were illustrated in Figure 2. The aim was for participants to have approximately 30-min sessions on five days each week. Sometimes they had two 30-min sessions in one day if they were coming to university on only four days a week. However, only2 participants out of 8, La and R made five sessions each week and the rest of the participants had between 2 to 5 sessions per week.

Learning to Accuracy Phase. The 60 flashcards were shuffled by the experimenter before the Learning to Accuracy Phase started; all 60 terms (without their definitions) were listed on the recording sheets (Appendix 5) in the order after they were shuffled. The sequence of presentation of the 60 cards was fixed as shown in Appendix 5 and was the same for every participant.

Participants were seated face to face with the experimenter across a table in the experiment room. The experimenter picked six cards for learning from the top of the pile of 60 to start with; and the rest of the cards to be learned (the To-Be-Learned Pile) were placed on the table on the left of the experimenter.

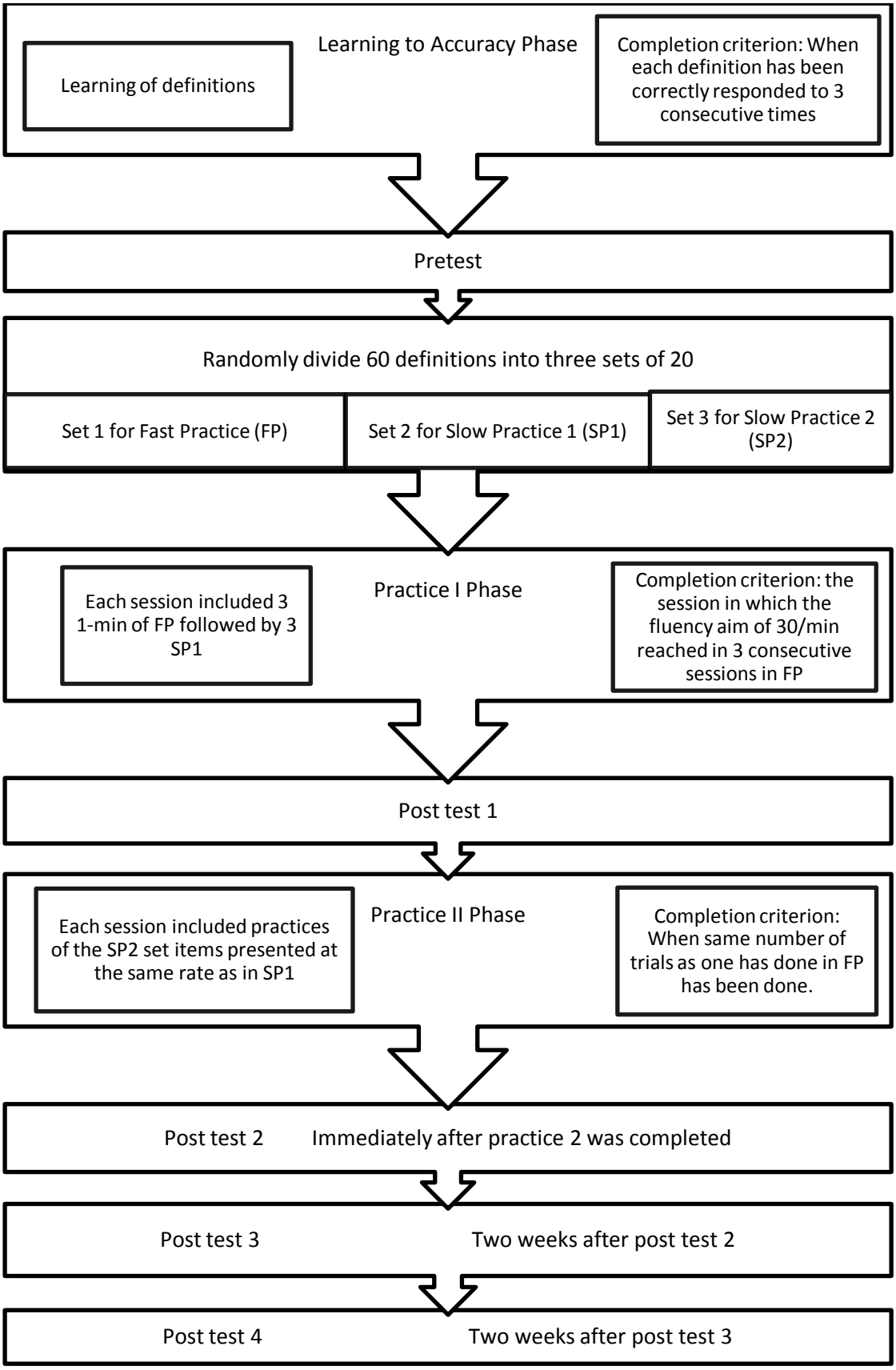


Figure 2. A diagrammed representation of the order of and procedures used in the study.

First, the experimenter held one of the six cards up, showing the side with a term on it; the participant was instructed to read it out loud. The experimenter then turned the card around and showed the back of it with its definition; and the participant was instructed to say it out loud. After the participant had read both sides of a card out loud, the experimenter then showed only the side of that card with the term on it. The participant was now instructed to read the term out loud and was given the opportunity to give the definition. Feedback like "good work/well done" was given for a correct response; and feedback as "close/try again" was given for an incorrect response and the term was re-presented until it was responded to correctly. The rest of the cards being learned were placed under the table so the participants were not able to see the next card before it was presented. The experimenter recorded on the recording sheet whether a response was correct or incorrect; or whether the participants simply read the definitions.

The initial six cards were shuffled after all of them had been correctly responded to once. Then participants were shown only the side of the card with the term on it and were given the opportunity to answer it without looking at the back of the card. A card was put at the bottom of the pile of the six when it was correctly responded to and replaced by another one from the top of the six. A card that had been correctly responded to in three consecutive trials was put aside as "learned" and replaced with a new card from the To-Be-Learned Pile. Learning to Accuracy Phase continued until each definition had been correctly responded to 3 consecutive times.

The 60 definitions were allocated to each of the practice set in the following procedure. The experimenter read the definitions (without the term) in a normal speech speed and it was recorded with the voice recorder on the computer. There were brief breaks between any two definitions to distinguish among definitions. After reading all of the definitions, the experimenter then played the voice recording and recorded on paper the

length of time between the beginning and the end of a definition. The definitions were then grouped together according to the range of the lengths of time taken to read each. Each group was shuffled first and divided into 3 piles. So at the end of this allocation procedure, there were three piles of 20 definitions with equal number of short, medium and long definitions within each. These three piles were then randomly allocated to the practice sets, one for Fast Practice (FP), one for Slow Practice 1 (SP1) and the third for Slow Practice 2 (SP2). Mean time taken to say all definitions was 4 s for the FP set, 4.5 s for the SP1 set and 4 s for the SP2 set. Appendix 4 gives the details of the items of the set they were allocated to.

Pre-test. Once all cards were learned to accuracy, there was a pre-test. Appendix 6 gives a PowerPoint presentation of the instructions and terms used for this. Participants completed the pre-test using the computer after the Learning to Accuracy Phase and right before they started the Practice I Phase. The computer displayed the instructions and presented only the term of each of the 60 items in a randomised order; and it recorded the response latencies for each item. Response latency referred to the length of time between the onset of presentation of one term and the key press that brought the presentation of the next term. The test began with the presentation of the first randomly selected term once the participants pressed the space bar to indicate they were ready to start. Participants then said out loud the definition of that term and pressed the space. The pre-test continued until all 60 items had been presented once. No feedback was given after each response. Participants were told to work at their own pace. The experimenter recorded whether each response was correct or incorrect on the recording sheet. Participants were told the number of correct and incorrect responses at the end of the test. Performance was video-recorded.

Practice I Phase I. After pre-test, the first practice phase started in the same session. Appendix 7 shows a Power point presentation of the instructions, sample terms and answers displayed by computer program during this phase.

Each session of Practice I Phase consisted of a period of three 1-min Fast Practice (FP) followed by a period of three un-timed Slow Practice (SP1). Items used in the Fast Practice were selected randomly by the computer without replacement from the FP set and those used in Slow Practice I were selected randomly without replacement from the SP1 set. A response made with all the words or with all the key words and no more than two non-key words missing was considered to be correct; and a response made with one or more than one key words missing or with all the key words and more than two non-key words missing was considered to be incorrect. If the complete definition was said but the space bar was hit before it was finished, the response was recorded as incorrect. The experimenter recorded whether the responses were correct or incorrect on the recording sheets for practice phases.

In the FP, participants were instructed to work as fast and accurately as they can for the three 1-min practices; and they were also told that their goal was getting 30 definitions correctly in 1 min. Once the participants pressed the space bar to indicate they were ready to start, the first term appeared on the computer screen. They had to say out loud its definition immediately and press the space bar to indicate they had done so. Once the space bar was hit, the answer for that term appeared on the screen below it. They pressed the space bar again when they had finished reviewing. The answer automatically disappeared and the blank screen appeared if participants had not pressed the space bar within 5 seconds and another space bar press brought the next term. In a small number of trials after 100% accuracy had been achieved and in the rate-building process, two of the participants pressed the space bar before they finished with the answer. These responses were recorded as incorrect and the participants were reminded at the end of that minute of FP that they had to finish the whole definition before they press the space bar; otherwise the responses would be counted as incorrect. The two participants made no more such mistakes.

The computer randomly selected a definition from the FP set without replacement for each trial; once all had been selected the computer started to randomly selecting from the whole FR set again until 1 min had finished**.** The rate of presentation in FP was controlled by the participant. The experimenter sat at a distance behind the participant but where she could clearly see what was on the screen so that she could determine whether the responses were correct or incorrect. There was a brief 30-s break between the 1-min practices. Participants were told the number correct and incorrect responses at the end of each 1-min practice.

Although the definitions (answers) appeared in every trial, there was no other feedback from the experimenter about whether a response was correct or incorrect during the 1-min practice. As more and more sessions were completed, participants became more familiar with the definitions. Once they got all 20 definitions correct, they started pressing the space bar very quickly to skip the review of the answers (definition) so that they could reach the goal of 30 correct per minute. However, for a small number of trials at the beginning when participants were building up their rate of responding, two of them, R and C, pressed the space bar to skip the answers (definitions) even when they had made an incorrect response, one with an error, say a key word missing or with the answer to a closely related term (e.g., Type I error for a Type II error). If this occurred again in a subsequent 1-min practice, they were reminded that they had responded incorrectly to that particular term at the end of that minute. This feedback was added to prevent participants from practicing the incorrect responses over and over again.

There was a 2-min break after FP was finished. In the first practice session, the experimenter gave instruction on how to graph the data. During the break, participants graphed the best result from the 3 practices for that session on the semi-log chart. Participants were praised (i.e., “well done! You get more correct in one minute!”) if there was an increase in the number of correct responses compared to the previous session; they were

praised for their effort of trying hard (i.e., “good try!”) and were encouraged to work faster next time if there was no increase.

SP1 started after the 2-min break. The instructions were similar to those in FP but participants were asked to work slowly and accurately (please refer to Appendix 6 for the instruction given); there was no rate aim for SP1. Once participants pressed the space bar to indicate they were ready to start, the first term from the SP1 set appeared on the screen. The procedure was the same as FP but the screen went blank for 3-s after every answer (definition) disappeared and a new term was automatically presented after 3 s had passed (whereas in FP, participants had to press the space bar to bring out the next term). There was a brief 30-s break between each slow practice. Participants were told the number of correct and incorrect responses at the end of the three blocks of SP1. The total number of trials for the SP1 period in each session was the same for La, R and C with that of the earlier FP period. At the end of the SP1 period, participants were told their best performance of the three as % correct and they graphed it on the accuracy graph. Participants were then praised if there was an increase in the percent correct shown on the graph (e.g. “well done! Getting more accurate!”); or they were encourage to get more accurate if there was no increase; or if 100% accuracy was obtained, they were praised and encouraged to keep the performance. Practice I Phase was completed when participants finished the session with the rate aim reached in the FP.

Post-test 1. Participants completed a post-test which was exactly the same as the pre-test in the same session Practice I Phase was completed.

Practice II Phase. Practice II Phase in which SP2 sets were used started the day after Practice I Phase was completed. Instructions and the rate of presentation of terms in SP2 were identical to SP1 and the total number of trials was the same as the total number of trials completed overall of FP. The only difference between SP1 and SP2 was that in SP2, the

number of sessions was determined by the participants. Participants were told the number of trials completed as session went on. Participants could also saw the number of trials completed and the total number of trials for SP2 on the top right corner of the start up screen of SP2. Practice II Phase was completed when the same number of trials as in FP/SP1 had been done.

Post-test 2. Participants completed another post-test (as for post-test 1) in the same session as the Practice II Phase was completed. Participants were told that they did not have to practice those terms any more for the experiment.

Post-test 3. Participants completed a post-test 2 weeks after post-test 2.

Post-test 4. Participants completed a post-test 4 weeks after post-test 2.

Results

Learning to Accuracy

Table 1 (A- G) shows for each item the responses given on each exposure to that item (response), the number of times each item was read, the number of times a response was correct, the number of times a response was incorrect and the total number of responses from the whole Learning to Accuracy Phase for the 6 (Table 1.A-F) participants who completed this phase and the 1 (Table 1.G) participant who nearly completed this phase.

The items are presented with the items grouped into the categories that they were allocated to for the practice phases (set 1=fast practice, set 2=slow practice 1, set 3=slow practice 2) rather than in the order they were presented to the participants, to give a clear presentation of differences between items in each practice set. Appendix 5 is the recording sheet used by the experimenter and it gives the order of actual presentation. Items are arranged in ascending order according to the total number of trials within each practice set, with the item with the fewest trials to accuracy first for each practice set. The total numbers of trials to accuracy for each item differed among items and across participants and items have been.

All items were read (*) on the first trial they were presented. After both the term and answer had been read once, the same term was presented again and the participant was given the opportunity to try answering it or to read it again. Correct responses are shown as + and incorrect responses as -. If they chose to read the definition again, it is recorded as *. An item was removed after 3 correct responses (i.e., +++). For example, in Table 1.A, the response list for item 1018 was *+-*+++. This means participant L read item 1018 once and then correctly answered it once. But she answered it incorrectly in a subsequent trial so she read item 1018 again (second * in the list). Item 1018 was removed when she answered it correctly on three consecutive trials (indicated by the last +++).

Table 1

Data from the learning phase for all participants, include item number, the condition each item was allocated and whether the response was correct (+) or incorrect (-). Trials that involved reading both sides of the card were indicated by an asterisk (*). Set 1 refered to fast practice (FP), Set 2 refered to slow practice 1 (SP1) and Set 3 refered to slow practice 2 (SP2).

A. Data for La

Item	Set	Responses (read"*"; correct "+", and incorrect "-")	No. read	No. correct	No. incorrect	Total
1001	1	* + + +	1	3	0	4
1002	1	* + + +	1	3	0	4
1004	1	* + + +	1	3	0	4
1028	1	* + + +	1	3	0	4
1032	1	* + + +	1	3	0	4
1041	1	* + + +	1	3	0	4
1053	1	* + + +	1	3	0	4
1055	1	* + + +	1	3	0	4
1068	1	* + + +	1	3	0	4
1072	1	* + + +	1	3	0	4
1081	1	* + + +	1	3	0	4
1082	1	* + + +	1	3	0	4
1058	1	* + - + + +	1	4	1	6
1063	1	* * - + + +	2	3	1	6
1077	1	* - * + + +	2	3	1	6
1018	1	* + - * + + +	2	4	1	7
1023	1	* + - * + + +	2	4	1	7
1059	1	* + - * + + +	2	4	1	7
1078	1	* + - * + + +	2	4	1	7
1076	1	* - * + - * * + + + +	4	5	2	11
1009	2	* + + +	1	3	0	4
1014	2	* + + +	1	3	0	4
1039	2	* + + +	1	3	0	4
1043	2	* + + +	1	3	0	4
1050	2	* + + +	1	3	0	4
1052	2	* + + +	1	3	0	4
1054	2	* + + +	1	3	0	4
1064	2	* + + +	1	3	0	4
1069	2	* + + +	1	3	0	4
1079	2	* + + +	1	3	0	4
1083	2	* + + +	1	3	0	4
1060	2	* - * + + +	2	3	1	6
1005	2	* + - * + + +	2	4	1	7
1037	2	* * - * + + +	3	3	1	7
1012	2	* - * - * + + +	3	3	2	8
1025	2	* + + - * + + +	2	5	1	8
1075	2	* + - + * + + +	2	5	1	8
1011	2	* + - * + - + + +	2	5	2	9
1033	2	* - * + - * + + +	3	4	2	9
1010	2	* - * + - * * + + +	4	4	2	10
1013	3	* + + +	1	3	0	4
1040	3	* + + +	1	3	0	4
1042	3	* + + +	1	3	0	4
1049	3	* + + +	1	3	0	4
1057	3	* + + +	1	3	0	4
1061	3	* + + +	1	3	0	4
1066	3	* + + +	1	3	0	4
1067	3	* + + +	1	3	0	4
1080	3	* + + +	1	3	0	4
1016	3	* * + + +	2	3	0	5
1048	3	* - * + + +	2	3	1	6
1070	3	* + - * + + +	2	4	1	7
1006	3	* - * - * + + +	3	3	2	8
1038	3	* - * + - * + + +	3	4	2	9
1062	3	* - * + - * + + +	3	4	2	9
1065	3	* - * + - * + + +	3	4	2	9
1015	3	* - * + + - * + + +	3	5	2	10
1071	3	* + - * + - * + + +	3	5	2	10
1056	3	* - * - * + - * + + +	4	4	3	11
1020	3	* - * + - * + - * + + +	4	5	3	12

Table 1 continued

B. Data for R.

Item	Set	Responses (read"*"; correct "+", and incorrect "-")	no. read	no. correct	no. incorrect	total
1002	1	* + + +	1	3	0	4
1004	1	* + + +	1	3	0	4
1018	1	* + + +	1	3	0	4
1023	1	* + + +	1	3	0	4
1028	1	* + + +	1	3	0	4
1053	1	* + + +	1	3	0	4
1055	1	* + + +	1	3	0	4
1068	1	* + + +	1	3	0	4
1081	1	* + + +	1	3	0	4
1032	1	* * + + +	2	3	0	5
1058	1	* + - + + +	1	4	1	6
1082	1	* + * + + +	2	4	0	6
1041	1	* + - * + + +	2	4	1	7
1063	1	* + - * + + +	2	4	1	7
1072	1	* - + * - + + +	2	4	2	8
1077	1	* - * + * + + +	3	4	1	8
1059	1	* + - * + * + + +	3	5	1	9
1078	1	* - * + - * + + +	3	4	2	9
1076	1	* - * - * - * + + +	4	3	3	10
1001	1	* + - * + - * + - * + + - * + + +	5	8	4	17
1005	2	* + + +	1	3	0	4
1009	2	* + + +	1	3	0	4
1043	2	* + + +	1	3	0	4
1050	2	* + + +	1	3	0	4
1060	2	* + + +	1	3	0	4
1064	2	* + + +	1	3	0	4
1083	2	* + + +	1	3	0	4
1011	2	* * + + +	2	3	0	5
1069	2	* + + + +	1	4	0	5
1010	2	* * * + + +	3	3	0	6
1039	2	* - * + + +	2	3	1	6
1052	2	* - * + + +	2	3	1	6
1014	2	* + - * + + +	2	4	1	7
1079	2	* + - * + + +	2	4	1	7
1012	2	* * + - * + - * + + +	4	5	2	11
1033	2	* * + * + + - * + + +	4	6	1	11
1037	2	* * * * + + - * + + + + +	5	7	1	13
1075	2	* + + - * + - * + + * + + +	4	8	2	14
1025	2	* + - * + + - * + + - * + - * + + +	5	9	4	18
1054	2	* - * - * - * + - * + * + + - * + + +	7	7	5	19
1042	3	* + + +	1	3	0	4
1049	3	* + + +	1	3	0	4
1066	3	* + + +	1	3	0	4
1057	3	* * + + +	2	3	0	5
1016	3	* + + + + +	1	5	0	6
1040	3	* + * + + +	2	4	0	6
1061	3	* - + + + +	1	4	1	6
1070	3	* + + + + +	1	5	0	6
1080	3	* + + + + +	1	5	0	6
1056	3	* + - * + + +	2	4	1	7
1065	3	* + - * + + +	2	4	1	7
1067	3	* + - * + + +	2	4	1	7
1006	3	* * * - * + + +	4	3	1	8
1020	3	* + + - * + + +	2	5	1	8
1038	3	* - * - * + - * + + +	4	4	3	11
1048	3	* - - * + + - * + + + +	3	6	2	11
1013	3	* + + - + - + + - + + +	1	8	3	12
1071	3	* + - * + - * + + + + +	3	7	2	12
1062	3	* + * + - * + + * + - * + + +	5	8	2	15
1015	3	* - + - - * + - + - * + - * + + +	4	7	6	17

Table 1. continued

C. Data for C.

Item	Set	Responses (read"*"; correct "+", and incorrect "-")	No. read	No. correct	No. incorrect	Total
1002	1	* + + +	1	3	0	4
1041	1	* + + +	1	3	0	4
1072	1	* + + +	1	3	0	4
1023	1	* - + + +	1	3	1	5
1055	1	* - + + +	1	3	1	5
1004	1	* + - + + +	1	4	1	6
1032	1	* - * + + +	2	3	1	6
1068	1	* * - * + + +	3	3	1	7
1077	1	* * - - * + + +	3	3	2	8
1078	1	* - * + - + + +	2	4	2	8
1082	1	* - * + - + + +	2	4	2	8
1018	1	* - * + - * + + +	3	4	2	9
1076	1	* - - * + + - + + +	2	5	3	10
1053	1	* - * + + - + - + + +	2	6	3	11
1058	1	* - * + + - * + + + +	3	6	2	11
1001	1	* - * + - * + + - * + +	4	5	3	12
1063	1	* - + + * + + - * + + +	3	7	2	12
1028	1	* + - * + + - * + - * + + + +	4	7	3	14
1081	1	* - - * + + - + + - + * + - * + + - * + + +	5	11	6	22
1059	1	* - * - * + * - * + + - * + * + - * + + * + + +	9	10	5	24
1050	2	* + + +	1	3	0	4
1069	2	* + + +	1	3	0	4
1079	2	* + + +	1	3	0	4
1083	2	* + + +	1	3	0	4
1005	2	* * + + +	2	3	0	5
1010	2	* - * + + +	2	3	1	6
1064	2	* - * + + +	2	3	1	6
1043	2	* + - * - + + +	2	4	2	8
1009	2	* - * + + * + + +	3	5	1	9
1052	2	* - * + - * + + +	3	4	2	9
1039	2	* * + - * - * + + +	4	4	2	10
1060	2	* - * + - * + + + +	3	5	2	10
1014	2	* - * + - * + + * + + +	4	6	2	12
1011	2	* - - * + - * + - + + +	3	5	5	13
1037	2	* - * - * - * + - * + + +	5	4	4	13
1075	2	* - * + - - * + + - * + + +	4	6	4	14
1012	2	* - * + - * + - * + - * + + +	5	6	4	15
1025	2	* + + - * + + - + - * + + - + + +	3	10	4	17
1054	2	* - * - * - * - + * + + * + + + + +	7	8	4	19
1033	2	* - * - * - * + - * + - * + - * + * + - + + +	8	8	7	23
1042	3	* + + +	1	3	0	4
1057	3	* + + +	1	3	0	4
1066	3	* + + + +	1	4	0	5
1016	3	* - * + + +	2	3	1	6
1040	3	* - * + + +	2	3	1	6
1049	3	* + - + + +	1	4	1	6
1061	3	* - * + + +	2	3	1	6
1080	3	* - * + + +	2	3	1	6
1006	3	* - * - * + + +	3	3	2	8
1056	3	* * - * - * + + +	4	3	2	9
1067	3	* - - * + - + + +	2	4	3	9
1070	3	* - * + - + + + +	2	5	2	9
1013	3	* - * + + - * + + +	3	5	2	10
1065	3	* - * + + - * + + +	3	5	2	10
1038	3	* - - * - - * + * + + +	4	4	4	12
1048	3	* - * + - * - - * + + +	4	4	4	12
1015	3	* - * - * + + - * + + + +	4	6	3	13
1020	3	* - * + - * + + - * + + +	4	6	3	13
1062	3	* - * - - * + - * + - * + + * + + +	6	6	5	17
1071	3	* * - * + - * + - * + - * + - * + - * + - * + + +	9	9	7	25

Table 1. continued

D. Data for T.

Item	Set	Responses (read "*"; correct "+", and incorrect "-")	No. read	No. correct	No. incorrect	Total
1002	1	* + + +	1	3	0	4
1004	1	* + + +	1	3	0	4
1058	1	* + + +	1	3	0	4
1068	1	* + + +	1	3	0	4
1032	1	* + - + + +	1	4	1	6
1041	1	* - * + + +	2	3	1	6
1063	1	* + * + + +	2	4	0	6
1072	1	* + * + + +	2	4	0	6
1078	1	* + + - * + + +	2	5	1	8
1053	1	* + - * + - * + + +	3	5	2	10
1055	1	* + + - * + - * + + +	3	6	2	11
1082	1	* + + - * + + - + + +	2	7	2	11
1023	1	* + - * + - + + - + + +	2	7	3	12
1081	1	* + - + - + - + * + + +	2	7	3	12
1059	1	* + - * + - * + - + + + +	3	7	3	13
1018	1	* + - * + + - * + - + + + + +	3	9	3	15
1077	1	* * - * + + * + + - * + + + +	5	8	2	15
1028	1	* + - * + - * + - * + + * + + +	5	8	3	16
1076	1	* * - * + * + - * + + - * + + +	6	7	3	16
1001	1	* + - * + - * + - * + - * + + * + + +	6	9	4	19
1043	2	* + + +	1	3	0	4
1050	2	* + + +	1	3	0	4
1064	2	* + + +	1	3	0	4
1009	2	* + - + + +	1	4	1	6
1069	2	* + - + + +	1	4	1	6
1005	2	* + + - * + + +	2	5	1	8
1079	2	* + + - * + + +	2	5	1	8
1083	2	* + - + * + + +	2	5	1	8
1010	2	* * + + - * + + +	3	5	1	9
1075	2	* + - * + - + + + +	2	6	2	10
1060	2	* - - * + * + + - + + +	3	6	2	11
1052	2	* - * + - * + - * + + +	4	5	3	12
1039	2	* * * * + - - + - + + + +	4	6	3	13
1014	2	* + - * + - + + - + + - + + +	2	9	4	15
1011	2	* * * + - + - + + + - + + + + +	3	10	3	16
1054	2	* - * + - * + - * + - * + + - * + + +	6	8	5	19
1037	2	* - - * + - * + + - + - * + + - * + - * + + +	6	10	6	22
1025	2	* + - + + - + - * + - * + + - * + * + + - + + +	5	13	6	24
1033	2	* * - * - * + - * + - * + + - * + + - * + * + + - * + + +	10	12	7	29
1012	2	* - * + - * + + - - * + - * + + - * + - + + - * + + - * + - + + +	8	15	10	33
1042	3	* + + +	1	3	0	4
1049	3	* + + +	1	3	0	4
1066	3	* + + +	1	3	0	4
1067	3	* + + +	1	3	0	4
1080	3	* + + +	1	3	0	4
1070	3	* * + + +	2	3	0	5
1016	3	* * - * + + +	3	3	1	7
1040	3	* - - * + + + +	2	4	2	8
1056	3	* - - * + + + +	2	4	2	8
1065	3	* - * + - + + +	2	4	2	8
1006	3	* - - * + - + + +	2	4	3	9
1057	3	* + + - * + + + +	2	6	1	9
1038	3	* - * + - * + + - * + + +	4	6	3	13
1013	3	* + - * + + - * + + - + * + + +	4	9	3	16
1062	3	* - + - + * + + - * + - * + + + + +	4	10	4	18
1061	3	* * - * + - * + - * + + - * + + - * + + +	7	9	4	20
1020	3	* - * + - * + - * + - * + + - * + + + + +	7	9	5	21
1015	3	* + - * + - * + * + + - * + - * + - * + + - * + + +	8	12	5	25
1071	3	* * + - * + - * + + - * + - * + - * + + - * + + +	8	11	6	25
1048	3	* - - * + - - * + - * + + - * + - + - * + * + + - * + + +	8	12	9	29

Table 1. continued

E. Data for A.

Item	Set	Responses (read"*"; correct "+", and incorrect "-")	No. read	No. correct	No. incorrect	Total
1004	1	* + + +	1	3	0	4
1053	1	* + + +	1	3	0	4
1058	1	* + + +	1	3	0	4
1068	1	* - * + + +	2	3	1	6
1072	1	* - * + + +	2	3	1	6
1077	1	* - * + + +	2	3	1	6
1002	1	* + + * + + +	2	5	0	7
1023	1	* + * - + + +	2	4	1	7
1078	1	* + - * - * + * + + +	4	5	2	11
1082	1	* - * + - * + + - + + +	3	6	3	12
1032	1	* - * - * + - * + - + + +	4	5	4	13
1063	1	* - - * - * + - * + - + + +	4	5	5	14
1076	1	* - - * - * + - * - - * + + +	5	4	6	15
1001	1	* + - * + + - * + - * + - + + +	4	8	4	16
1041	1	* + - + + - - * + - + - + * + + +	3	9	5	17
1081	1	* - * - * + - * + - * - * + + - + + +	6	7	6	19
1018	1	* + - * + - + - * + + - + - + + - + * + + +	4	12	6	22
1028	1	* + - * + - + + - * + - * + + - + + * + - + + +	5	13	6	24
1055	1	* - * - * + - * + - * + - * + + - * + - * + - * + + - * + + +	11	13	10	34
1059	1	* - * + - + - - * + + - * + - * + + - * + - * + - * + - * + + - * + - * + - + - * + - + * + + +	13	20	15	48
1050	2	* + + +	1	3	0	4
1069	2	* + + +	1	3	0	4
1064	2	* * + + +	2	3	0	5
1010	2	* * - * + + +	3	3	1	7
1043	2	* + - + + * + + +	2	6	1	9
1052	2	* * + - * - * + + +	4	4	2	10
1083	2	* + + - * + + * + + +	3	7	1	11
1005	2	* - * + - * + + - + + +	3	6	3	12
1079	2	* - * + - * + + - + + +	3	6	3	12
1054	2	* - * + + - - - - * + + +	3	5	5	13
1060	2	* - * - * - + - + + - + - + + +	3	7	6	16
1009	2	* + - * + * + + - - * + + * + + +	5	9	3	17
1011	2	* - * - * + - + + * + - * + + - + + +	5	9	5	19
1039	2	* - * - * - * + - * + - * + + - - * + + +	7	7	7	21
1075	2	* - * + - * + - * - * + - * + + * + + - + + +	7	10	6	23
1025	2	* - - * + - + - * + - * + - + + - + - + - * + +	5	12	9	26
1012	2	* - - * + - * + - * + - - * + - + - * + - * + - * + - - * + + +	9	11	12	32
1014	2	* - * + - + + - * + + * + - * + - * + - + * + + - + + + + + + +	7	18	7	32
1033	2	+ + * - * + + - * + - * + - * - * + - * - * + - * + - * + - * + - * + - * + * + + +	14	14	12	40
1037	2	* + - * - * - * - * + - - * - * - * + - - - * + - * + - - * + + - * + - * + - * + - * + - * + - * + - + + +	16	15	20	51
1080	3	* + + + + +	1	5	0	6
1066	3	* - * + * + + +	3	4	1	8
1042	3	* + * + + - + + +	2	6	1	9
1049	3	+ + + * + - * + + - + + +	2	6	2	10
1040	3	* - * + - * - * + + +	4	4	3	11
1057	3	* + + - + + * + + + +	2	8	1	11
1006	3	* - * - * + - * + * + + +	5	5	3	13
1065	3	* - + - - * + - + - + + +	2	6	5	13
1067	3	* + - * - * + - * + + - * + + +	5	7	4	16
1056	3	* - - * + - + - + - - * + + - + + +	3	8	7	18
1070	3	* + - + + * + - + - * + - + + + + +	3	11	4	18
1038	3	* * - * - * - * - * + - * + - * + + +	9	5	7	21
1015	3	* - * + + - - * + - + - * + - - * + * + + +	6	9	7	22
1016	3	* - - * - * + - * + + - + - * + - + * + + +	6	9	7	22
1048	3	* - * + - * + - - - * + - * + + - + - + - * + - + + +	6	11	10	27
1061	3	* + - * + - * + - * + - * + + - * + - * + - * + - * + - * + * + + +	11	14	9	34
1013	3	* - - * + - * + - * + - * + + - * + + * + - + + * + - + - * + - + + +	10	17	11	38
1062	3	* - * + - * - * + - + - * + - * + - + - - * + - * - - * + - + - * + + - * + - - * + - * + - * + * + + +	15	18	18	51
1071	3	* - * - * + - * - * + - - * + - * - * + - * + - * + - - * + - * + - * - + - * + - * + * + - * + - + + + +	17	17	19	53
1020	3	* - * + * + + - + - * + - * + - * + - * + * + - * + - * + - * + + - + * + - + - + - * + + * + - + + - - * + + +	15	25	16	56

Table 1. continued

F. Data for L.

Item	Set	Responses (read"*"; correct "+", and incorrect "-")	No. read	No. correct	No. incorrect	Total
1002	1	* + + +	1	3	0	4
1004	1	* + + +	1	3	0	4
1018	1	* + + +	1	3	0	4
1041	1	* + + +	1	3	0	4
1053	1	* + + +	1	3	0	4
1055	1	* + + +	1	3	0	4
1072	1	* + + +	1	3	0	4
1082	1	* + + +	1	3	0	4
1063	1	* - * + + +	2	3	1	6
1068	1	* + * + + +	2	4	0	6
1078	1	* + * + + +	2	4	0	6
1023	1	* + + * + + +	2	5	0	7
1032	1	* - + + * + + +	2	5	1	8
1058	1	* + + - * + + +	2	5	1	8
1076	1	* - * + - * + + +	3	4	2	9
1081	1	* - + * + - + + +	2	5	2	9
1001	1	* + - * + * + - + + +	3	6	2	11
1077	1	* - + + - * + + - + + +	2	7	3	12
1059	1	* + + - * + - * + + * + - * + + +	6	10	4	20
1028	1	* + - * + - * + - * + - * + - * + - * + + +	7	9	6	22
1005	2	* + + +	1	3	0	4
1009	2	* + + +	1	3	0	4
1050	2	* + + +	1	3	0	4
1064	2	* + + +	1	3	0	4
1083	2	* + + +	1	3	0	4
1010	2	* * + + +	2	3	0	5
1012	2	* - * + + +	2	3	1	6
1060	2	* - * + + +	2	3	1	6
1069	2	* + + + + +	1	5	0	6
1079	2	* - * + + +	2	3	1	6
1011	2	* + - * + + +	2	4	1	7
1039	2	* - * + + * + + +	3	5	1	9
1075	2	* + - + - + - + + +	1	6	3	10
1052	2	* * + - * + - * + + +	4	5	2	11
1054	2	* - * + - * + * + + +	4	5	2	11
1014	2	* + - * + * + - * + + +	4	6	2	12
1043	2	* + - * + - * + * + + +	4	6	2	12
1025	2	* + + - * + - * + + - + + - + + +	3	10	4	17
1037	2	* - * + - * + - * + - + - + - * + - * + - + - - * + * + + +	8	12	10	30
1033	2	* - + - * + + - * + - * + - * + * + + - * + + - * + + - * + + - * + + +	9	15	8	32
1040	3	* + + +	1	3	0	4
1042	3	* + + +	1	3	0	4
1049	3	* + + + +	1	4	0	5
1062	3	* - + + +	1	3	1	5
1067	3	* - * + + +	2	3	1	6
1057	3	* + + * + + +	2	5	0	7
1066	3	* + - * + + +	2	4	1	7
1006	3	* - * + - * + + +	3	4	2	9
1080	3	* + + * + + + + +	2	7	0	9
1061	3	* - * - + + * + + + +	3	6	2	11
1071	3	* + - * + + - * + + +	3	6	2	11
1048	3	* - * + - * + - * + + +	4	5	3	12
1070	3	* + - * + + - + * + + +	3	7	2	12
1016	3	* + - * - * + + - * + + +	4	6	3	13
1056	3	* + - * + * + - * + - * + + + +	5	8	3	16
1065	3	* - * + - * + - * + * + - * + + + + +	6	9	4	19
1038	3	* - * - * - * - * - * - * - * - * + - * + - + + - * + - * + - + + +	13	9	15	37
1020	3	* + - * + - * + - * + - * + * + - * + - * + - * + - * + + * + - * + - * + + +	13	16	10	39
1013	3	* - + - * + - * + - * + - * + * + - * + - * + - * + - * + - * + - * + - * + + + +	14	16	13	43
1015	3	* - * - * - * - * - * - * - * + - * - * - * + - * - * + - * + - * + - * + + - * - * + - * - * + + +	20	11	19	50

Table 1. continued

G. Data for J.

Item	Set	Resonses (read"*"; correct "+", and incorrect "-")	No. read	No. correct	No. incorrect	Total
1018	1		0	0	0	0
1077	1		0	0	0	0
1053	1	* + + +	1	3	0	4
1082	1	* + + +	1	3	0	4
1068	1	* * + + +	2	3	0	5
1002	1	* + - + + +	1	4	1	6
1023	1	* + - + + +	1	4	1	6
1032	1	* + * * + + +	3	4	0	7
1004	1	* + - + - + + +	1	5	2	8
1058	1	* * + - * + + +	3	4	1	8
1055	1	* + + * - * + + +	3	5	1	9
1063	1	* - * - - * + + +	3	3	3	9
1041	1	* - * + * - * + + +	4	4	2	10
1072	1	* - * + * - - + + +	3	4	3	10
1081	1	* * - * - * + + * + + +	5	5	2	12
1078	1	* - - - * + - - - * + - * + + +	4	5	7	16
1076	1	* - * + + - * + - * + + * - + + +	5	8	4	17
1059	1	* - * + - + * - * + - * + + * + - * + + +	7	9	5	21
1028	1	* + - * + * + - * + * + - * + - + * + + - + + * + - + + +	8	15	6	29
1001	1	* - * + + - * + - * - * + + * + - * + - - * - - * + * + - + + - + + * + - * + + +	13	17	14	44
1069	2		0	0	0	0
1075	2		0	0	0	0
1050	2	* + + +	1	3	0	4
1083	2	* + + +	1	3	0	4
1043	2	* - + + +	1	3	1	5
1064	2	* - + + +	1	3	1	5
1011	2	* + + * + + +	2	5	0	7
1009	2	* + - + - * + + +	2	5	2	9
1010	2	* - * - * - + + +	3	3	3	9
1014	2	* - * - * - + + +	4	3	3	10
1039	2	* + + * - - * + + +	3	5	2	10
1052	2	* + * - * - - + + +	3	4	3	10
1079	2	* + * - * - * + + +	4	4	2	10
1025	2	* + * + - - - * + + +	3	5	3	11
1060	2	* - * + + * + - + + +	3	6	2	11
1005	2	* - + - * + + - * + + +	3	6	3	12
1054	2	* * * + + * - - * + + +	5	5	2	12
1012	2	* - * - * + + * - * + - - * + + +	6	6	5	17
1033	2	* - * + - * + - + - + - + * - * + + * + + +	6	10	6	22
1037	2	* - * - * + - + - * + - * - + - * - * + * - + - + * + - - * + - - + * + + +	11	13	14	38
1006	3		0	0	0	0
1056	3		0	0	0	0
1065	3		0	0	0	0
1067	3		0	0	0	0
1070	3		0	0	0	0
1080	3		0	0	0	0
1042	3	* + + +	1	3	0	4
1049	3	* + + +	1	3	0	4
1066	3	* * + + +	2	3	0	5
1040	3	* - + - + * - - + + +	2	5	4	11
1013	3	* - * + - * + - * + + +	4	5	3	12
1057	3	* + - * + * + + - + * + + +	4	8	2	14
1016	3	* - * + + - * + - * + - * + + +	5	7	4	16
1020	3	* - * + - * + - * + - * + + * + + + +	7	10	5	22
1015	3	* - * - * - * - * + - * - * + - * + + * + + +	9	7	7	23
1048	3	* + - * + + * + - - * + * + - - + - + * + + +	6	11	6	23
1061	3	* - * + - * + + - * + + * + + - * + + * + + +	7	12	4	23
1038	3	* * * - * + - + + - * + + - * + - * + - * + + * + + +	9	12	6	27
1062	3	* - * - * + - - * + - * - - * + - + * - - * + - + * + + * + - + + +	10	12	12	34
1071	3	* - * - * + - - + - + * + - - + * + - * - * - + - + * + - + - * - + * + - + + +	10	15	15	40

Note. Blank lines indicating item not completed in learning phase.

Some items have more than 3 + at the end of their response list, such as item 1076 in Table 1.A, item 1069, 1037, 1016 in Table 1.B. This is because such items were used to make up the total of 6 cards when the number of cards not learned was fewer than 6. Moreover, some items had * (read) right after a + (correct response) as a result of the item being correctly responded to at the end of one session but not having reached criterion in that session. Thus the next session started with a "read" for that item. Examples of these items are 1075 in Table 1.A, 1075, 1054 in Table 1.B, and 1062 in Table 1.C.

As can be seen in Table 1 (A-G), the number of trials with each item varied. For example, for participant La (Table 1. A), the number of trials for items 1001, 1002, 1004 was 4 (the minimum), and the number was 11 for item 1076 and 8 for item 1012. Distributions of items with fewer or more trials appear to be somewhat uneven across practice sets. There were more items with fewer trials in the FP set and the SP1 set than in the SP2 set for participant La (Table 1. A) and R (Table 1. B); more items with fewer trials in the FP set than in the SP1 and SP2 sets for Participant La (Table 1. F) and Participant A (Table 1. E). There were similar numbers of items with fewer trials across the 3 practice sets for Participant C (Table 1. C), T (Table 1. D) and J (Table 1.G). The pattern of differences in numbers of items with larger number of trials relative to the other items across practice sets was similar among participants. There were more items with larger numbers of trials for the SP1 and SP2 practice sets for participants La, R, T, Le and A; there were more in SP2 than in FP and SP1 for Participant J; and there were more in SP1 for than in FP and SP2 Participant C.

The total number of trials for the items differed across participants. For example, while the numbers of trials range from 4 to 12 for participant La (Table 1.A), the range for participant R (Table 1.B) was 4 to 19, and it was 4 to 25 for participant C (Table 1.C) and 4 to 33 for participant T (Table 1.D) and so on. Generally it was different items that gave the

different participants more trouble (more trials), but some items appear to have been difficult for several participants (e.g. 1059 and 1076 in set 1, 1033 in set 2, and 1071 in set 3).

Interobserver Agreement

Interobserver agreement was calculated for approximately 50% of the recorded sessions, 1/6 of the total number of trials occurred (3976). A recent graduate student in Masters of Applied Behavior Analysis was first trained by the experimenter to score whether a response was correct or incorrect. He then watched the video and scored the responses independently. The interobserver agreement was calculated by dividing the total number of agreed and disagreed instances by the number of agreed instances and multiplied by 100 (Cooper, Heron, & Heward, 2007). The interobserver agreement score obtained was 99.7% which was considered to be high according to Cooper et al. (2007).

Practice

The number of graphed sessions and trials over the practice phases is shown in Table 2 and it differed across participants. There were some sessions not graphed on the chart and Table 2 only shows the number of sessions with graphed data. The number of trials reflects the total number of trials. For the three participants who completed who completed all phases there were between 21 and 64 graphed sessions corresponding from 1248 to 3671 trials. For the two who did not completed, there were 11 and 19 graphed sessions and 362 and 362 and 713 trials respectively completed.

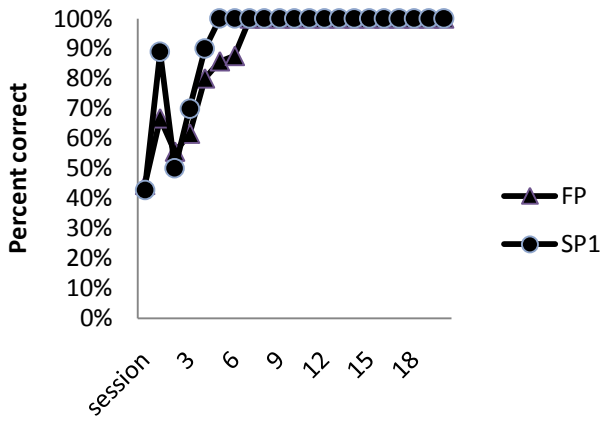
Accuracy. Figure 3 (A-E) (Also in Appendix 10 which shows the actual graphed used by the participants) shows the percentage of correct responses over the graphed sessions in Practice I Phase for FP and SP1 set for the five participants who started the practice phase.

Table 2

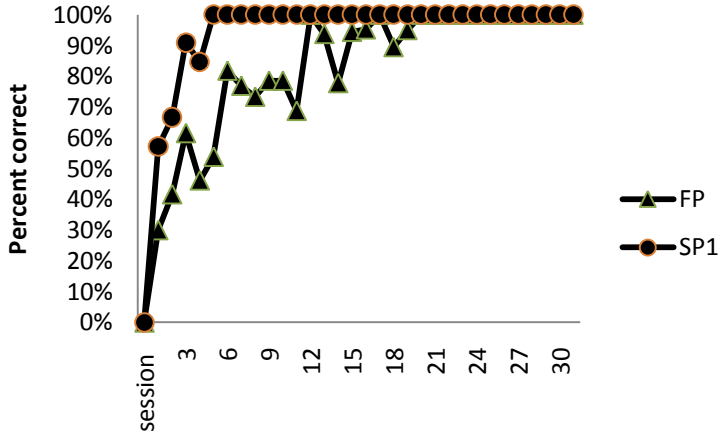
Total number of sessions and trials for each practice set for all participants in the Practice Phase I & II. Participant A did not finish SP1 after she had completed the FP in two sessions thus the total number of trials was fewer in SP1 than in FP for her.

Participant	FP & SP1		SP2	
	No. of sessions	No. of trials	No. of sessions	No. of trials
La	21	1248	N/A	1248
R	31	1959	N/A	1959
C	64	3671	N/A	3671
T	11	362		
A	19	713/651		

A. Data for La



B. Data for R.



C. Data for C.

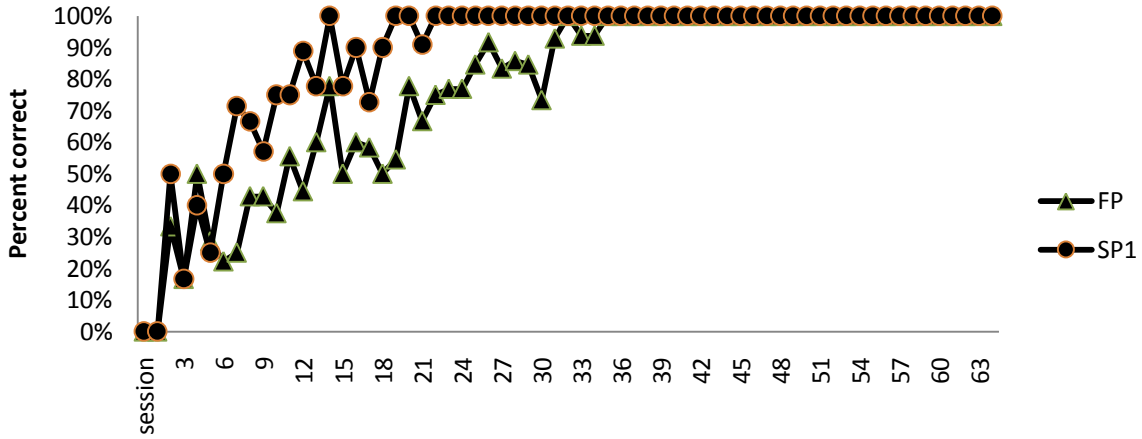
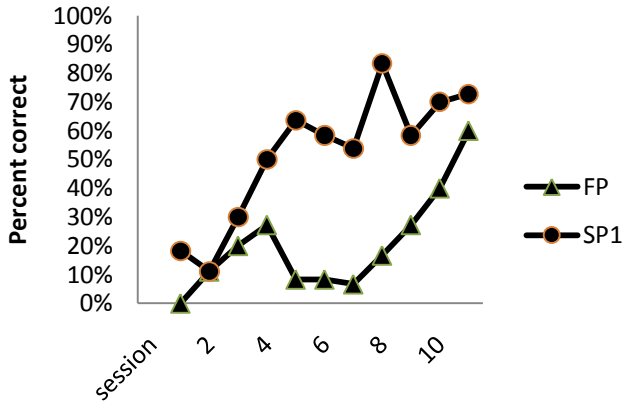


Figure 3. The percent of correct responses for each graphed sessions for FP and SP1 sets during Practice I Phase for participant La, R and C.

D. Data for T



E. Data for A.

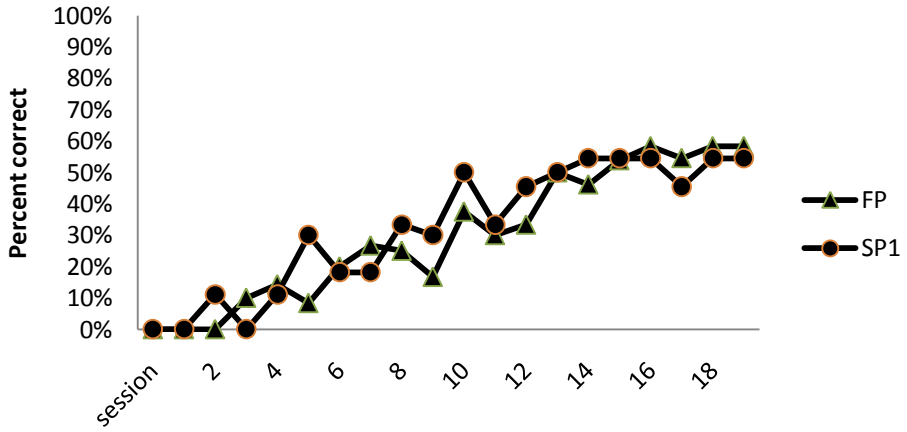


Figure 3 continued. The percent of correct responses for each graphed sessions for FP and SP1 sets during Practice I Phase for participant T and A.

The number of trials often differed between the best result of the FP and SP1 in a session, so although the number of correct responses was recorded for the fast practice, percent correct was used for comparison for accuracy. As can be seen in the figure, accuracy improved with more sessions completed and the rate of improvement was faster for SP1 than for FP for participants R, C, and T; and the improvement was similar between the two practice sets for participants La and A. For the three participants, La, R and C, who completed Practice I Phase, accuracy reached 100% with fewer practice sessions for SP1 than for FP. Participants La, R and C reached 100% accuracy for both practice sets on the 7th, 20th

and 34th graphed sessions. Participants T and A did not reached 100% accuracy for either of the practice sets at the time they withdrew from the study.

Celeration charts. Figure 4 (A-E) showed the semi-log charts downloaded from the website AimChart (2009) on which the experimenter entered the data for the fast practice collected from the experiment. See Appendix 9 for the actual semi-log charts used in the experiment. Charts for all the five participants who started the practice phase were presented. As can been seen in the charts, the number of correct responses per minute increased over practice sessions for all participants; and three of them, La, R and C, reached the rate aim of 30 correct per minute over 21, 31 and 64 sessions respectively. It appears that it took participants quite some sessions to reach the rate aim after they had achieved 100% accuracy. The number of graphed sessions required to reach the rate aim after 100% accuracy for La, R and C was 14, 11 and 30 respectively.

Response latencies (RL). Response latencies (RL) for each item over all trials were plotted for each practice set in Figure 5 (A-F); the vertical lines mark the end of all trials for each item. The latency to complete a response decreased over trials with longer latencies at the start of practices for all of the participants. For participants La, R and C, RLs for the majority of the items decreased from the beginning to the end of the practices. Exceptions for a small number of items existed where RLs stayed the same from the beginning till the end of the practices. The FP items tended to have shorter latencies than SP1 or SP2; the latencies for SP1 did not differ a lot from SP2. The graphs suggested that the distribution of response latencies tended to be skewed by some very long latencies. Thus, the median response latency may reflect the distribution better than the mean. However, it was decided in further analysis to present both the mean and the median.

Figure 4.A. Semi-log Chart for participant La. Score 1 in the graph shows when participants reaching the rate aim of 30/min.

40

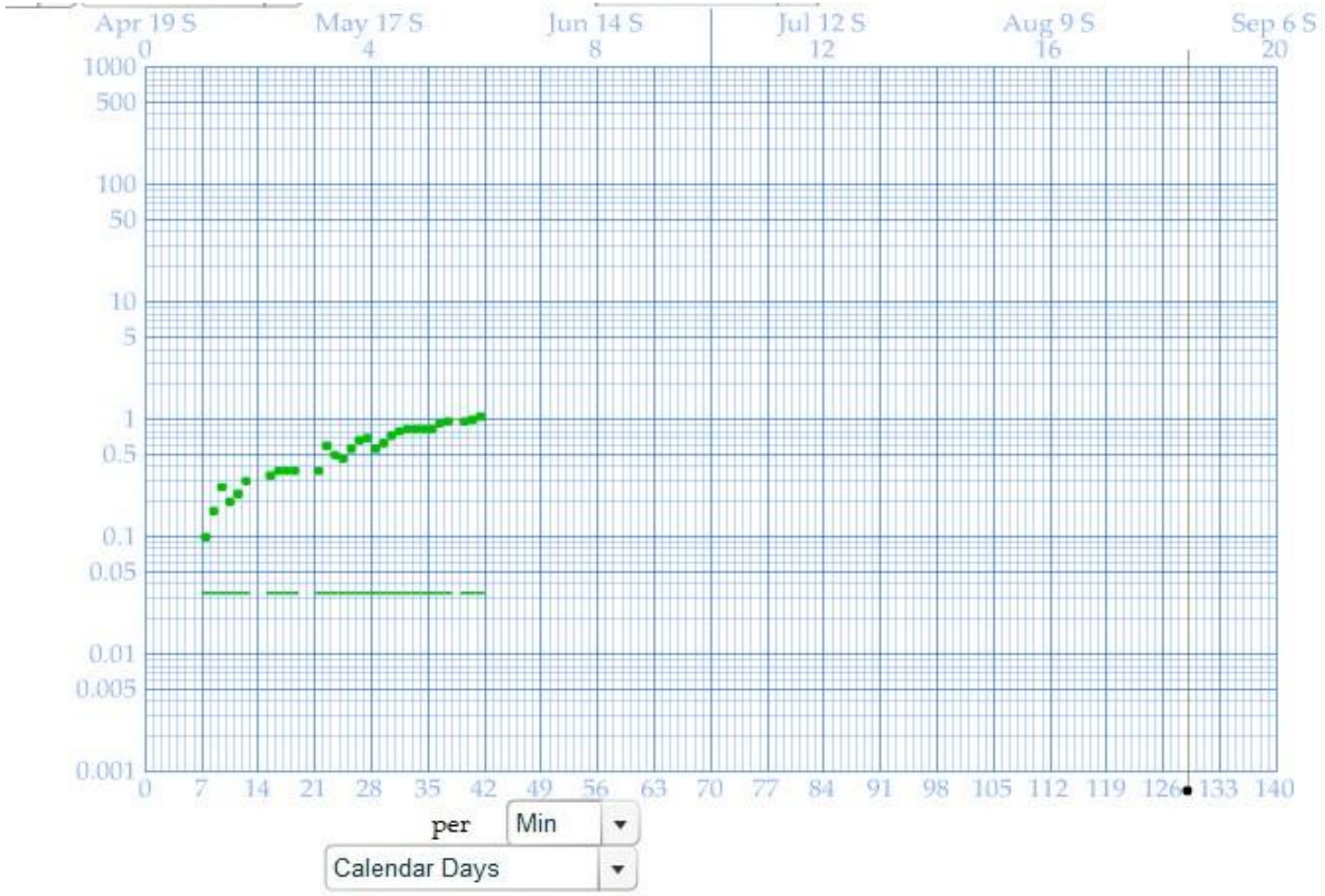


Figure 4.B. Semi-log Chart for participant R. Score 1 in the graph shows when participants reaching the rate aim of 30/min.

41

*Figure 4.*C. Semi-log Chart for participant C. Score 1 in the graph shows when participants reaching the rate aim of 30/min.

42

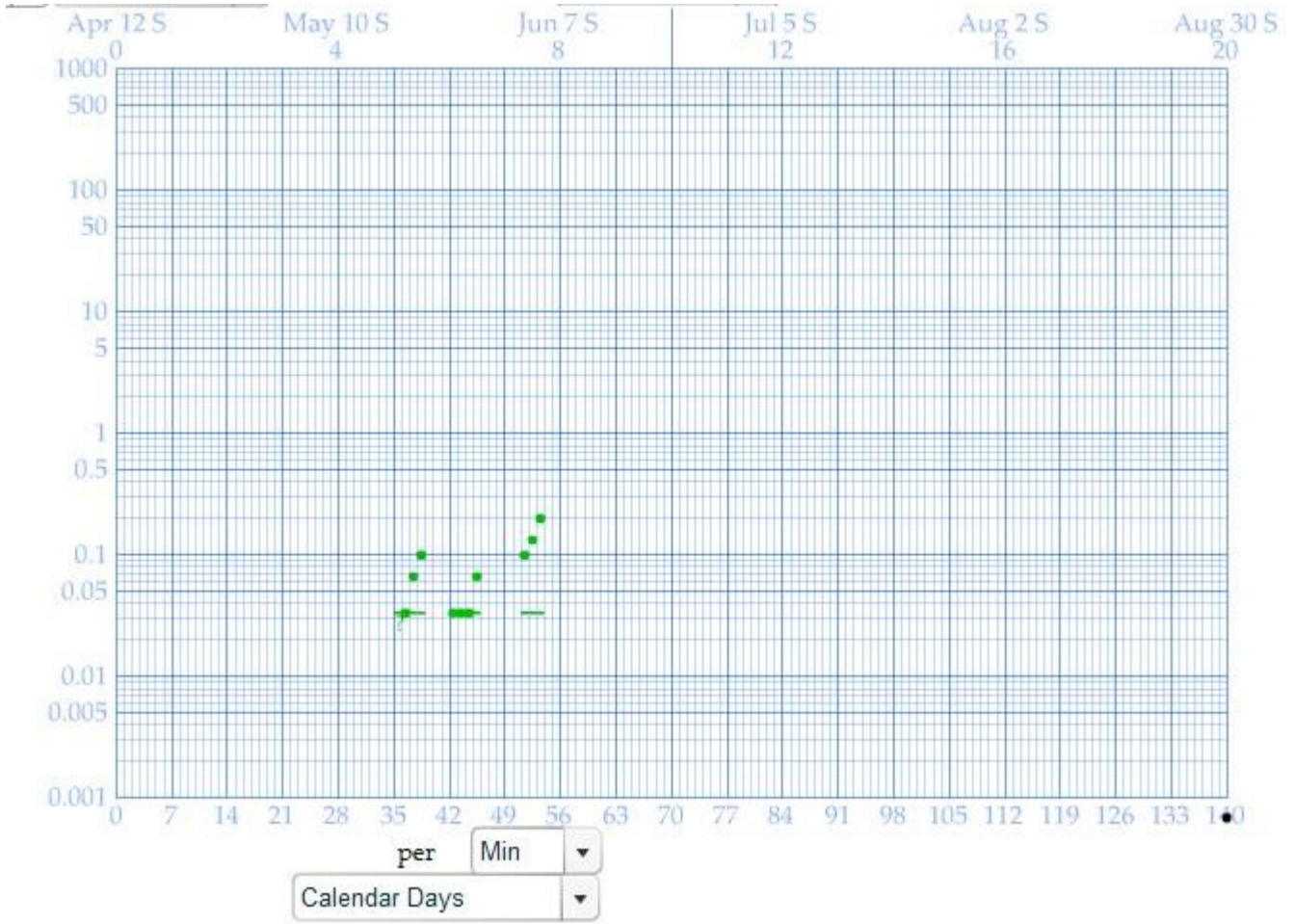


Figure 4.D. Semi-log Chart for participant T. Score 1 in the graph shows when participants reaching the rate aim of 30/min.

43

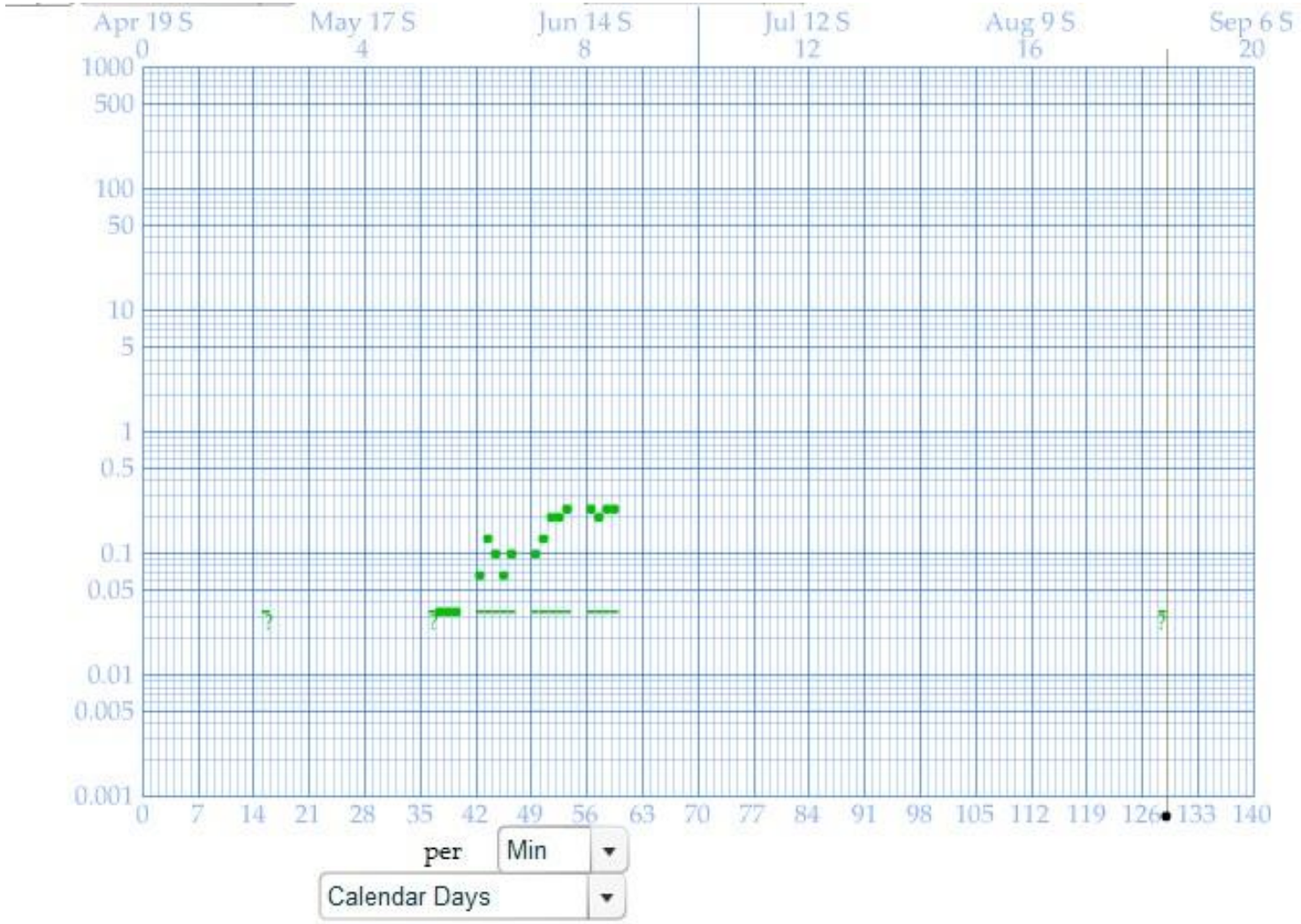


*Figure 4.*E. Semi-log Chart for participant A. Score 1 in the graph shows when participants reaching the rate aim of 30/min.

44

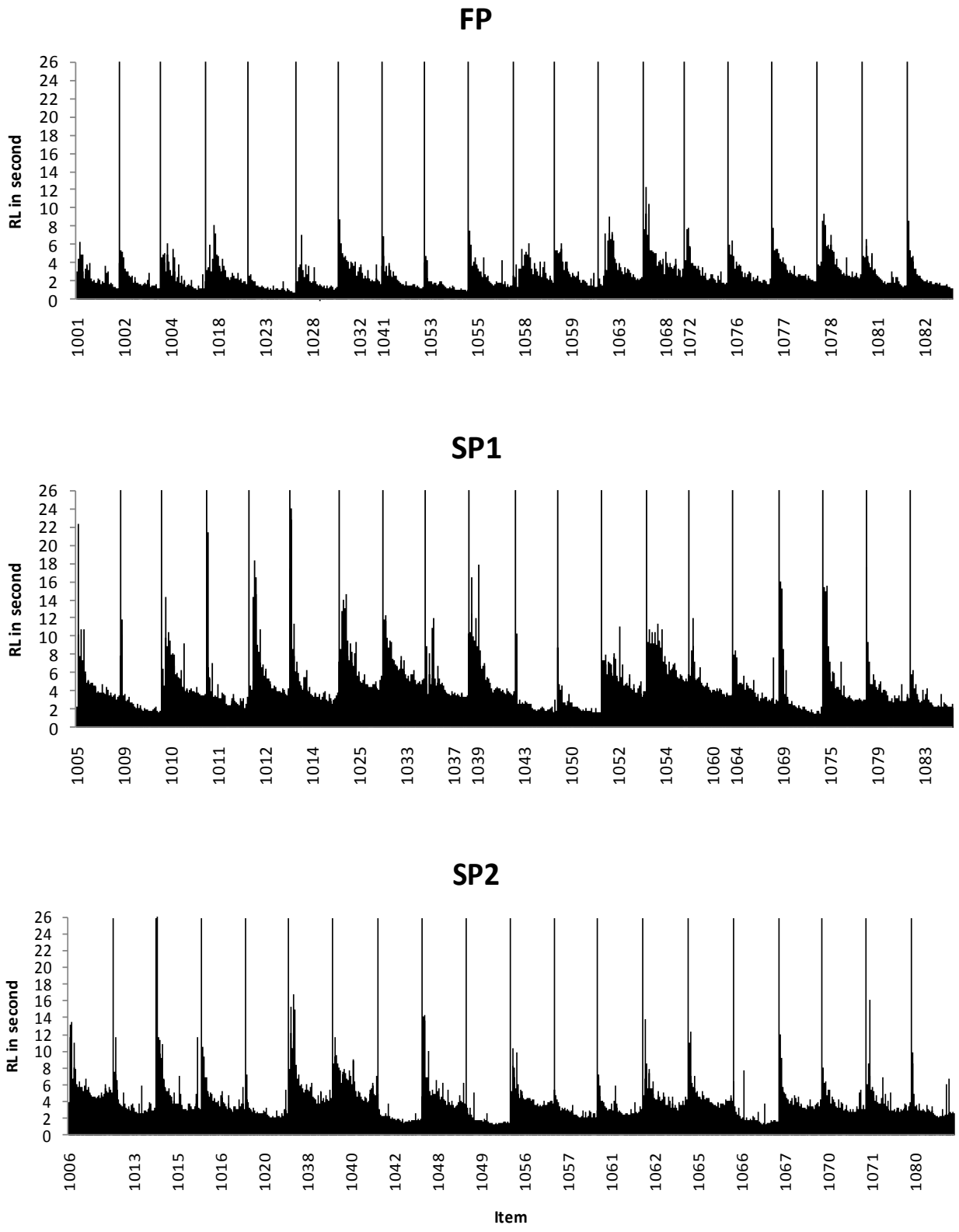


*Figure 5.*A. Response latencies for each item over the practice period for participant La.

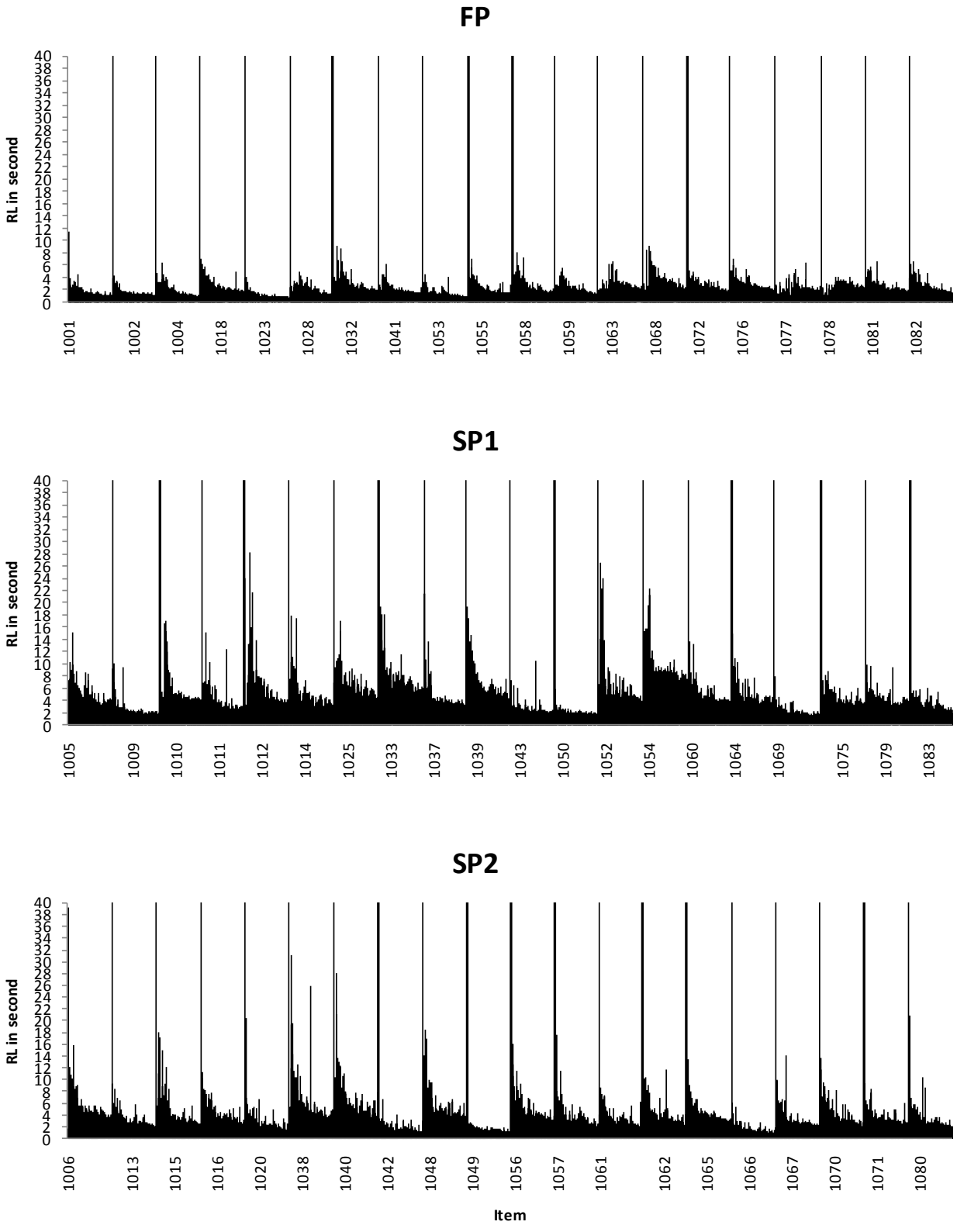


Figure 5.B. Response latencies for each item over the practice period for participant R.

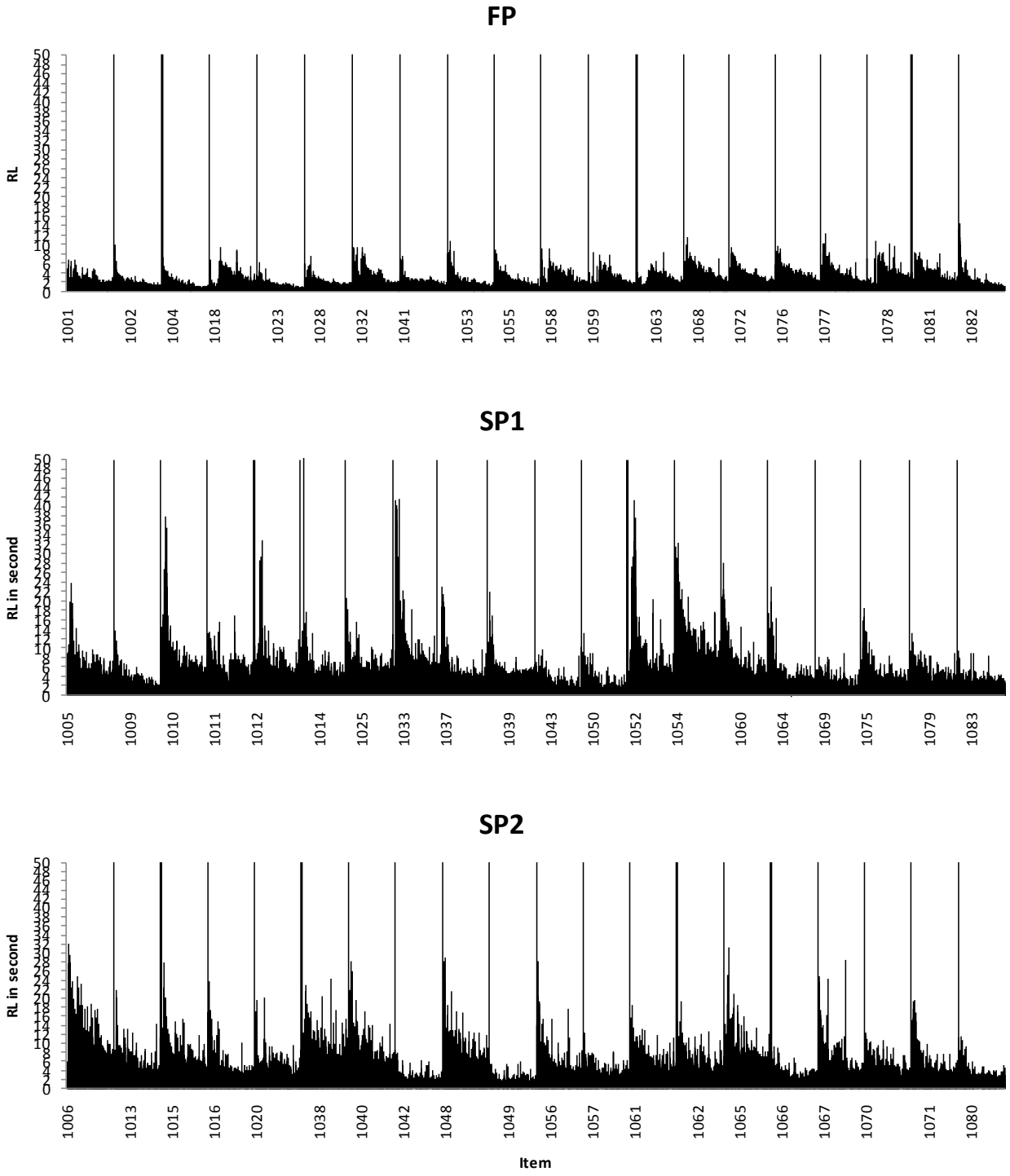


Figure 5.C. Response latencies for each item over the practice period for participant C.

Evans and Evans (1985) suggested that comparing the data of response rates at the beginning sessions and at the finishing sessions could reveal increase in rate of responding (if there was any) as the result of practice. Thus, the medians of the early 20 sessions and the final 20 sessions were used in current study examine whether there were changes in the latency for completing a response to an item from the beginning to the end of the practice phase.

The mean and median RLs of the first and last 20 sessions for each practice set are shown in Figure 6 to allow comparisons. Mean RLs were generally higher than median RLs in all practices for all participants; however, both showed the same pattern. Mean and median RLs of the first 20 sessions were longer than those of the last 20 sessions in all practice sets for all participants. For the 3 participants who completed all the phases, mean and median RLs of the last 20 sessions were shorter in FP than in SP1; and they were shorter in FP than in SP2. For participant La and C, mean and median RLs of the last 20 sessions were similar between SP1 and SP2; whereas for Participant R, they were shorter in SP2 than in SP1.

Test

Although there were test data for 5 of the participants who started the practice phases, the following section will deal mainly with data of the 3 participants who completed all phases and all tests.

Accuracy. The number of correct responses for each practice set in the pre-test and in the four post-tests is plotted in Figure 7. As shown in the figure, accuracy for participants La, R and C was low in pre-test but increased to 100% (e.g. 20 correct responses out of 20 for SP1 and SP2 for La and R) or almost 100% (e.g. 19 correct responses out of 20 for FP for La and SP1 and SP2 for C) soon after practice (ie. FP and SP1 in post-test 1 and SP2 in post-test

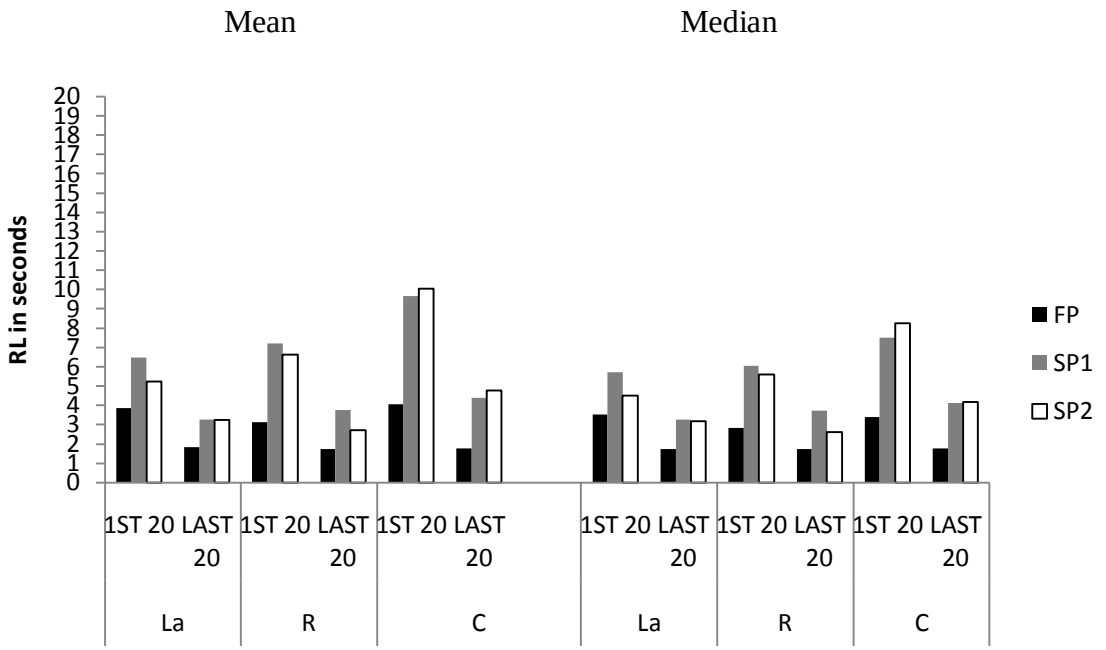


Figure 6. Mean and median response latencies (RL) for the first and last 20 sessions for participants La, R and C.

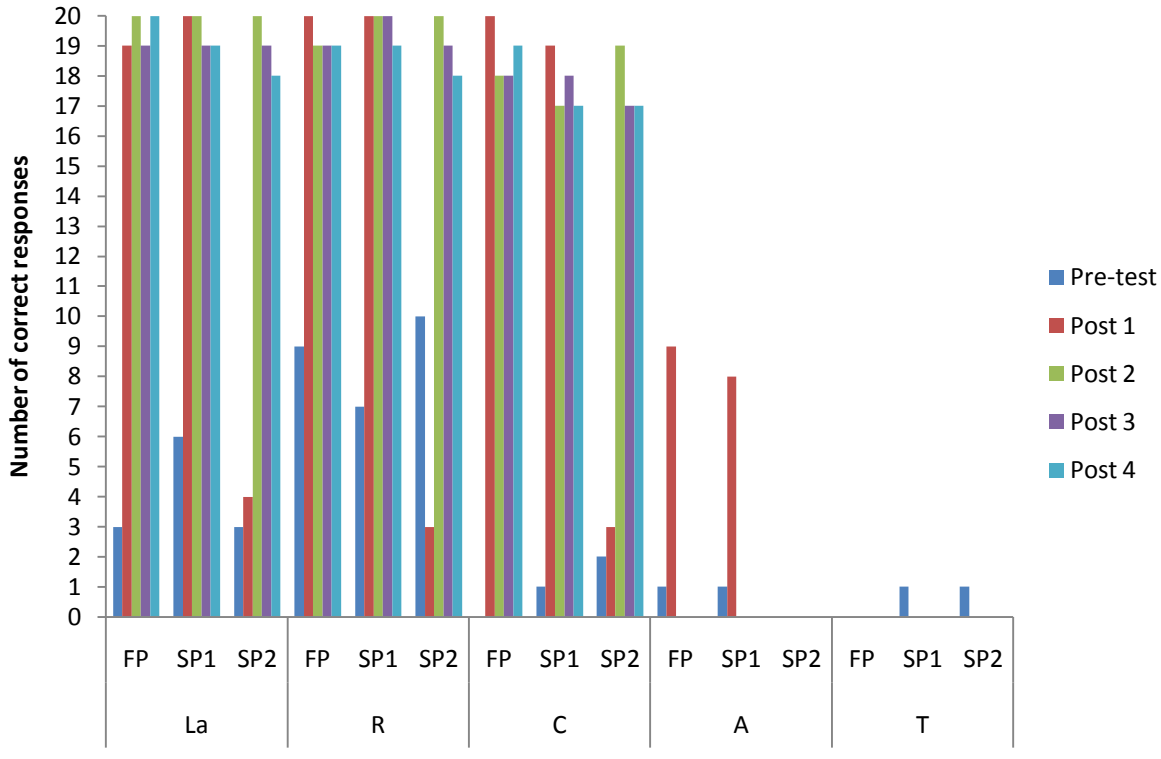


Figure 7. The number of correct responses for each practice set in the tests for participants La, R, C, T and A. *Note.* Participants C and T scored 0 for FP in the pre-test and participant A scored 0 for SP2 in the pre-test and post-test 1; and a score of 0.1 was entered to show that there was data (of 0) for that test.

2). Participant A had not reached the rate aim by the time she withdrew from the study but there were similar improvements for FP and SP1 in accuracy from pre-test to post-test 1.

Although participants differed in overall performance, for example, participant C had fewer correct responses in SP1 and SP2 in all post tests compared to La and R, there was no difference in accuracy between practice sets; the largest number of errors in one practice set after practice was only 3 out of 20 (as in SP1 and SP2 for participant C.).

Response latencies (RL). Response latencies for each item in each practice set over the pre-test and the 4 post-tests for the 3 participants who completed the study are shown in Figure 8. For comparison, data for all participants are plotted on the same scaled axes. RLs decreased from pre-test to post-test 1 and 2 then increased again in post-test 3 and 4 for most of the items in the 3 practice sets. Furthermore, some items had extremely long RLs (20 s, 58 s, and 65 s etc.) while some had very short ones (about 2 s) within each practice set. Most of the longer RLs are in the pre-test for La but for R and C there were some very long latencies in the post-tests in all practice sets. For example, as shown in Figure 8.B., the RL for item 1063 (FP) in post-test 2 was about 54 s; it was about 30 s for item 1064 (SP2) in post-test 4; and it was about 54 s for item 1062 (SP2) in post-test 4.

Figure 9 shows mean and median RLs over all items in each practice set in all tests along with the mean and median RLs of the last 20 sessions for comparison. As seen in the graphs, there were greater discrepancies between tests using the mean RLs compared to median RLs for all practice sets. All mean and median RLs in post-tests were longer than those in the last 20 sessions for each practice set. Mean and median

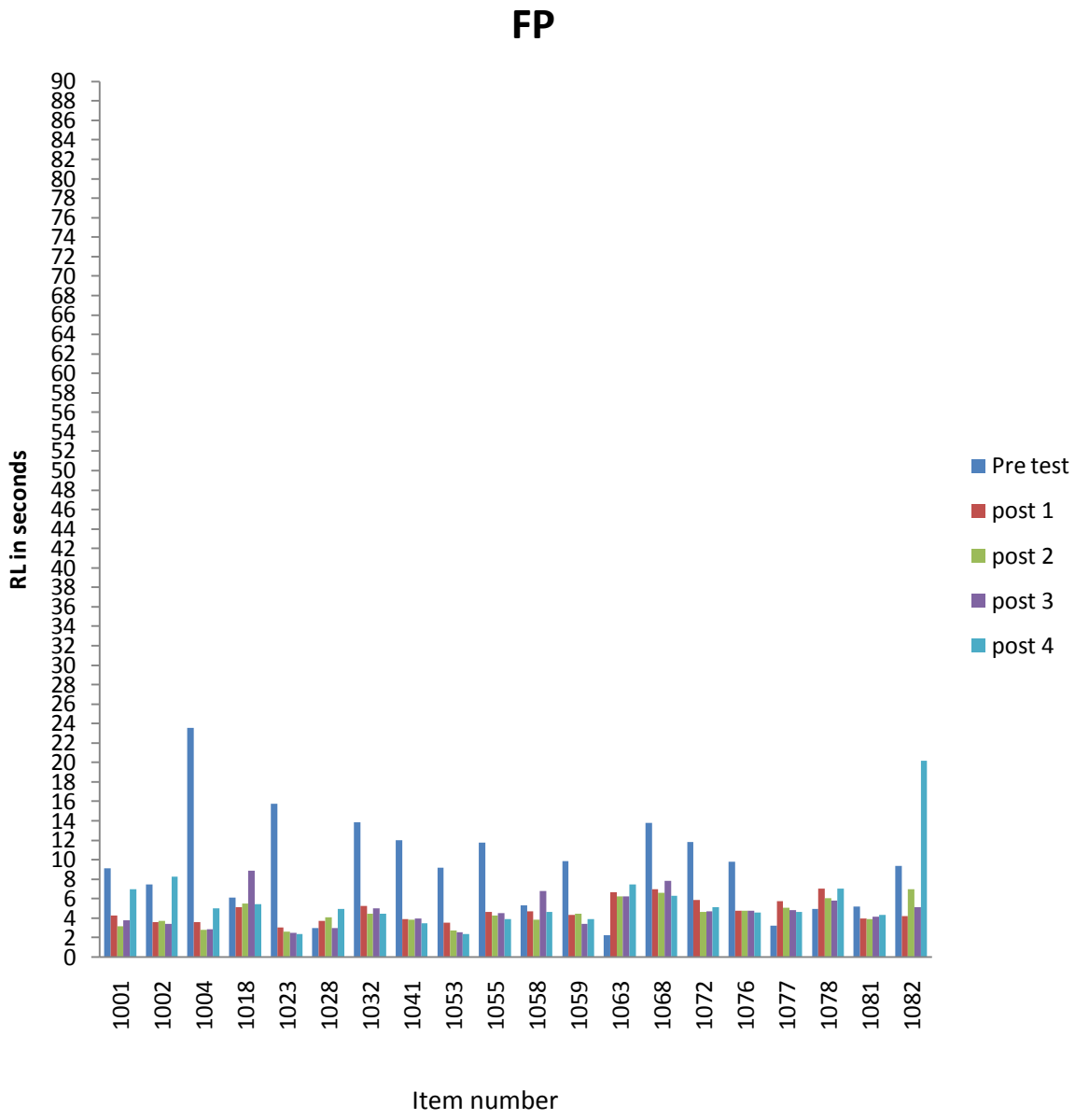


Figure 8.A. Response latencies for each item of the FP set in all tests for participant La.

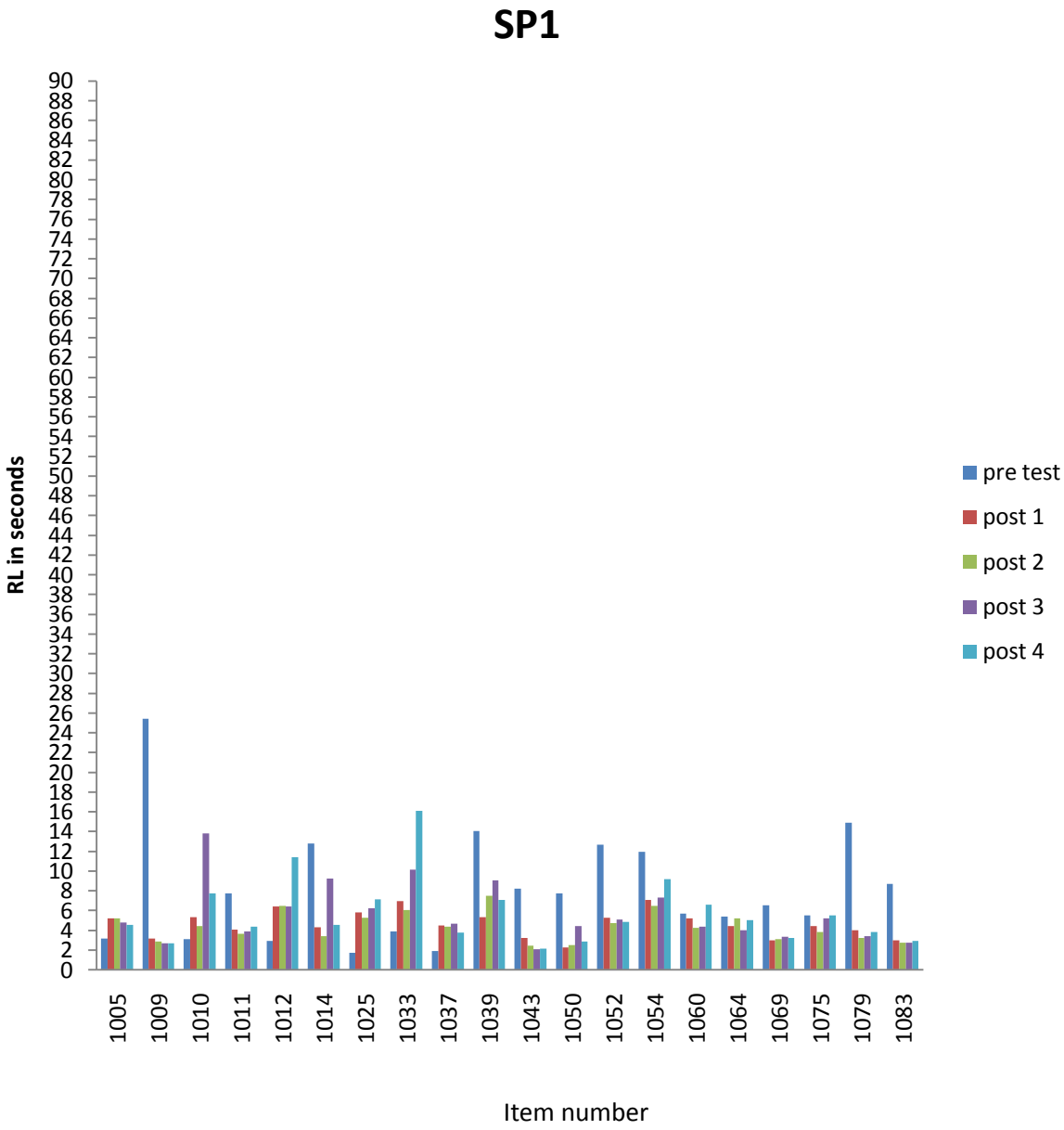


Figure 8.A.continued. Response latencies for each item of the SP1 set in all tests for participant La.

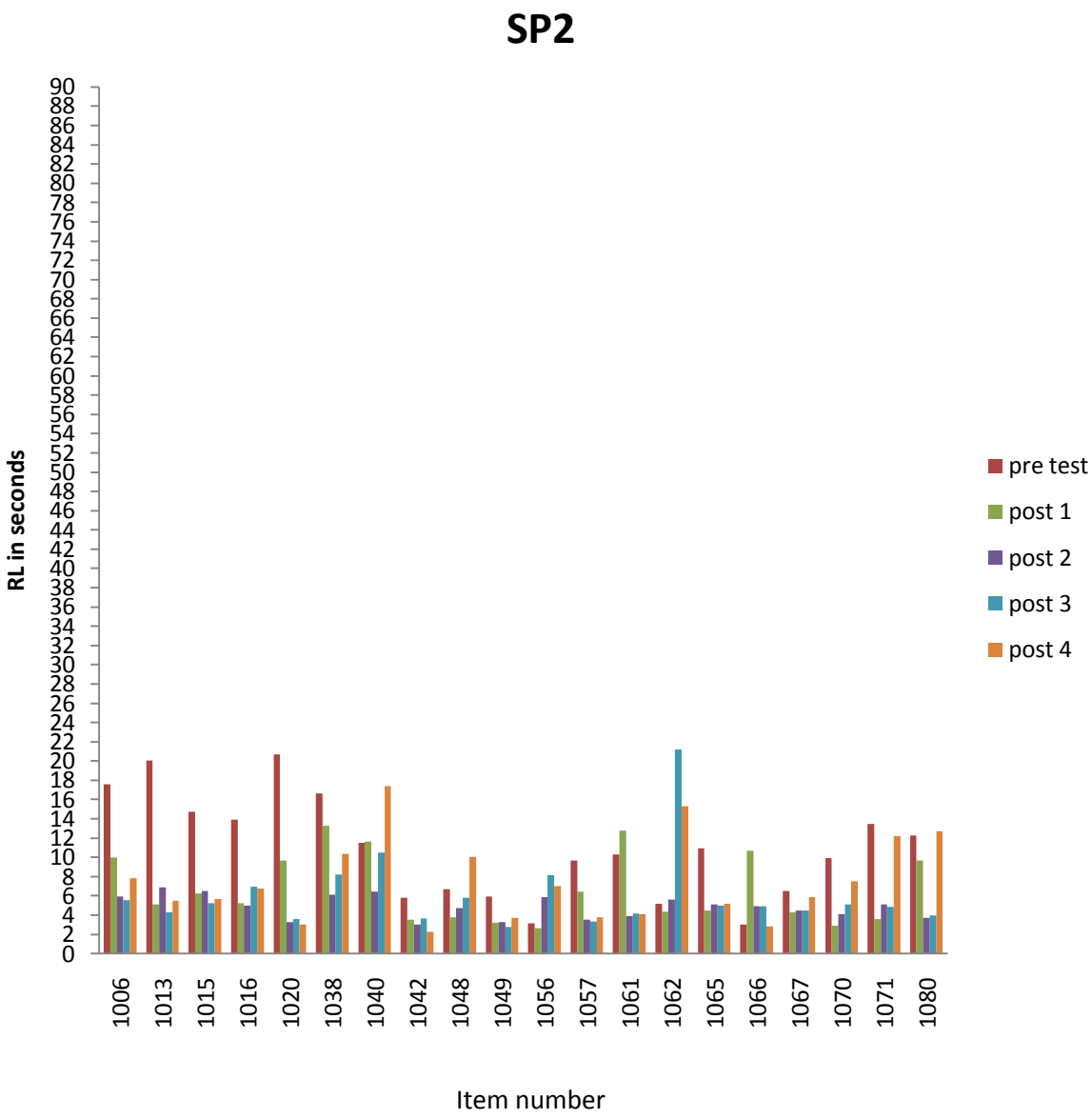


Figure 8.A.continued. Response latency for each item of the SP2 set in all tests for participant La.

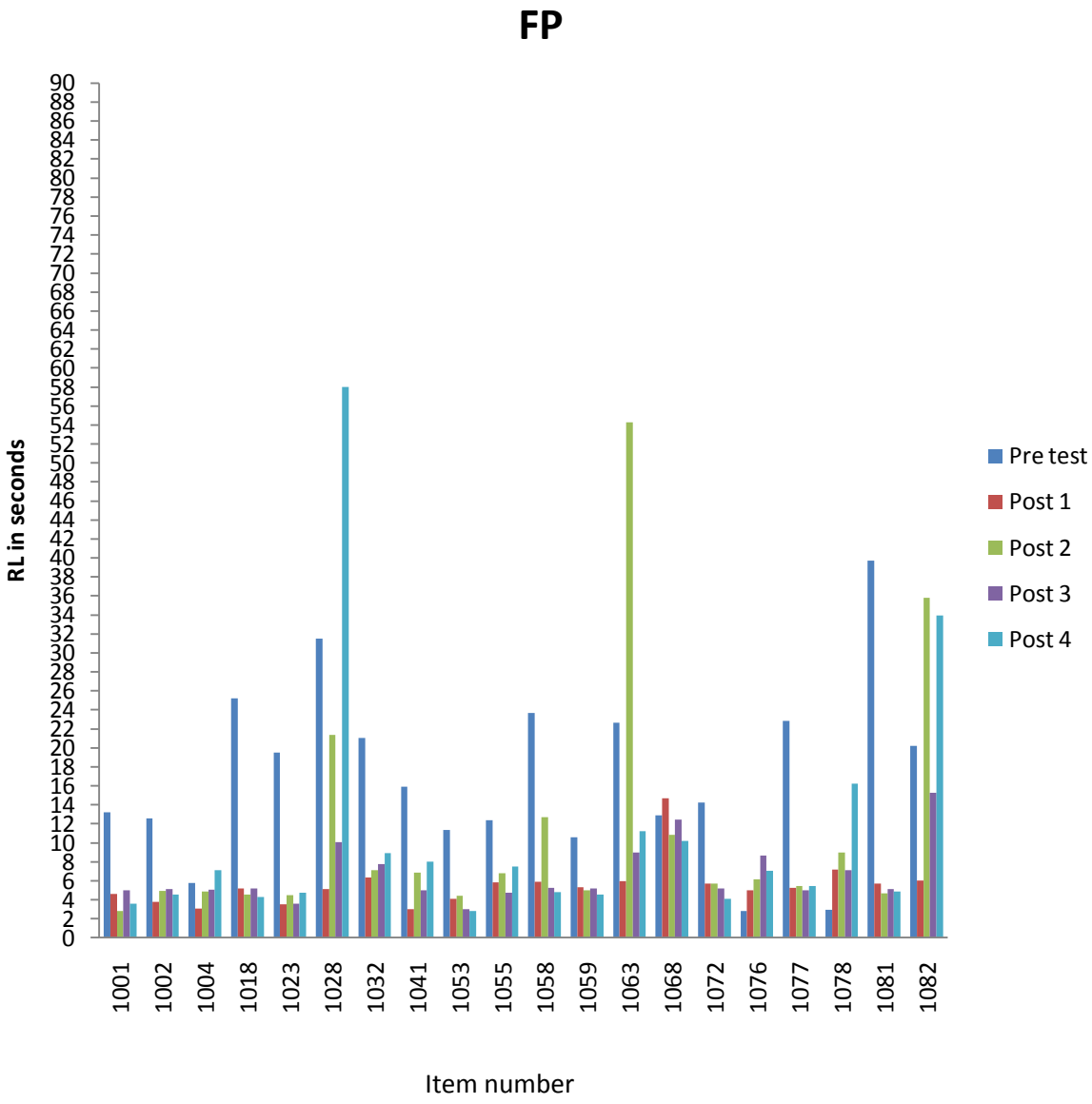


Figure 8.B. Response latency for each item of the FP set in all tests for participant R.

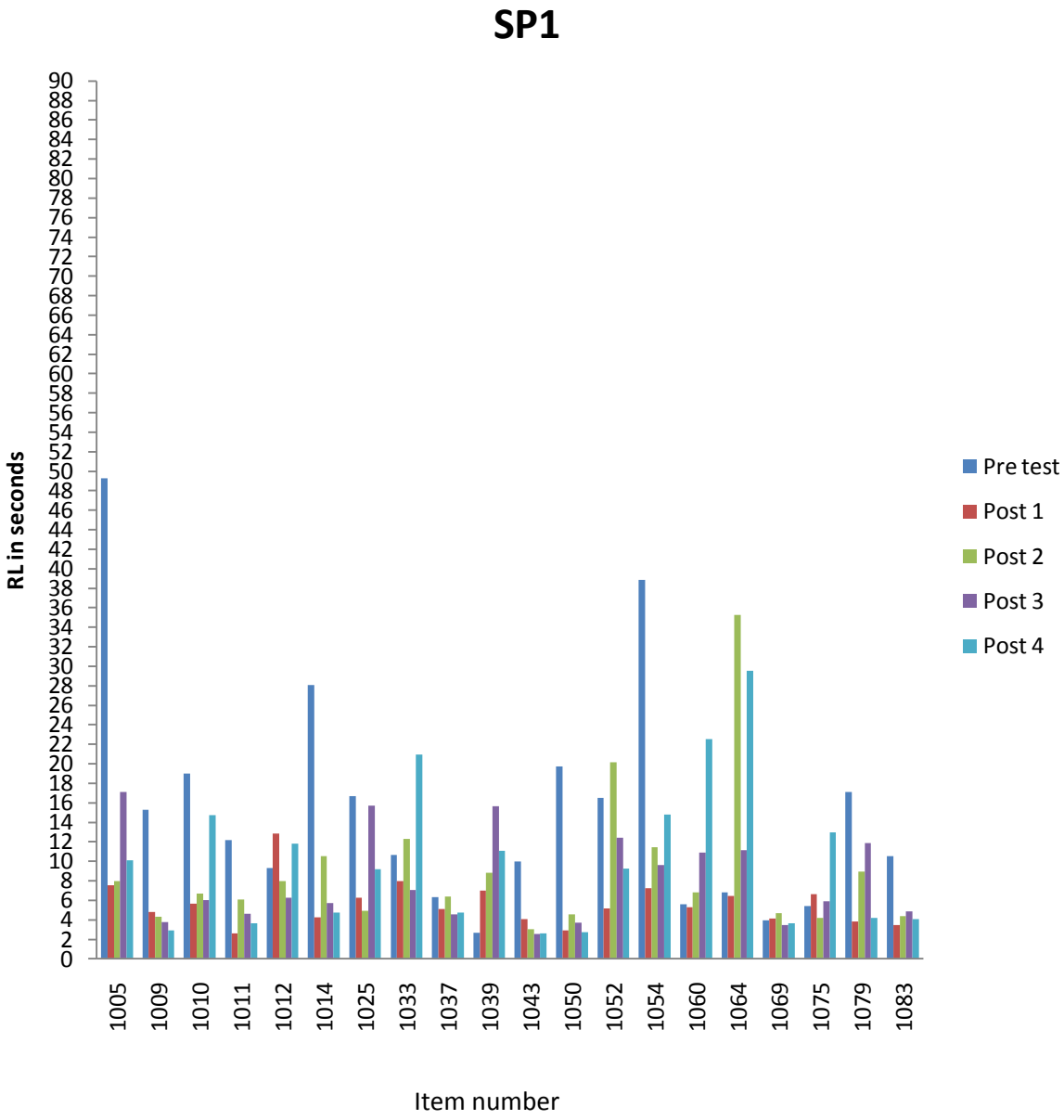


Figure 8.B.continued. Response latency for each item of the SP1 set in all tests for participant R.

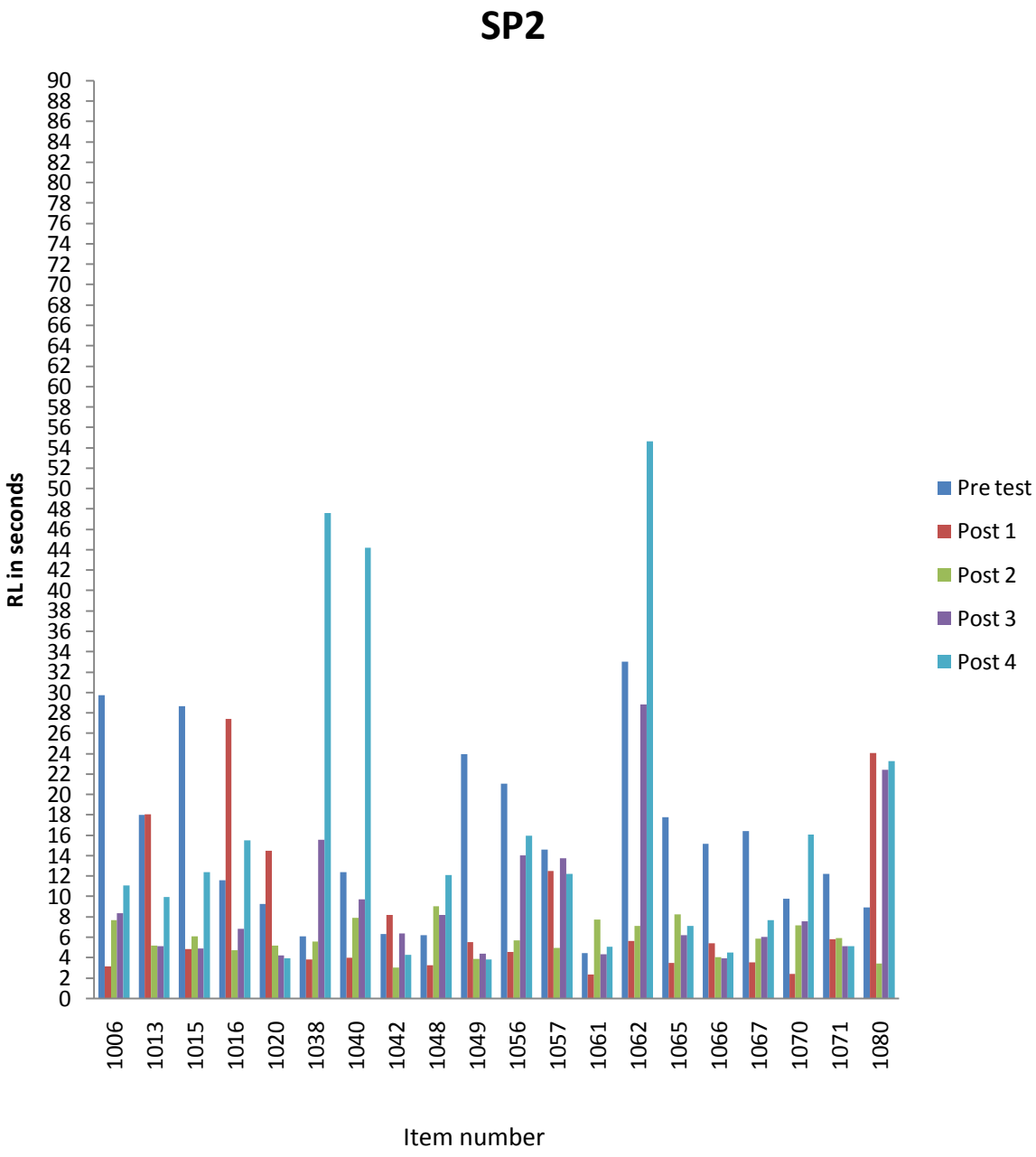


Figure 8.B.continued. Response latency for each item of the SP2 set in all tests for participant R.

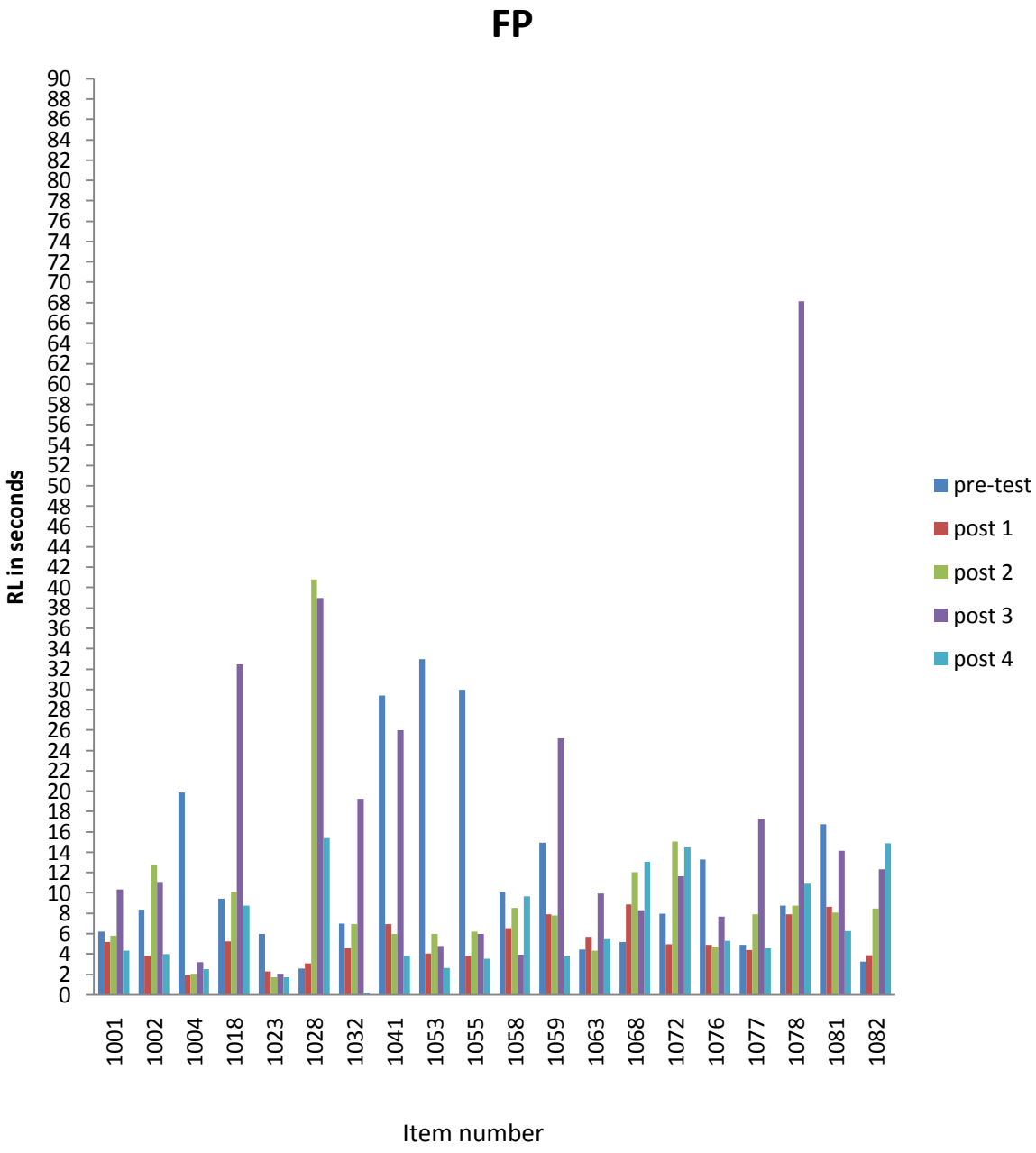


Figure 8.C. Response latency for each item of the FP set in all tests for participant C.

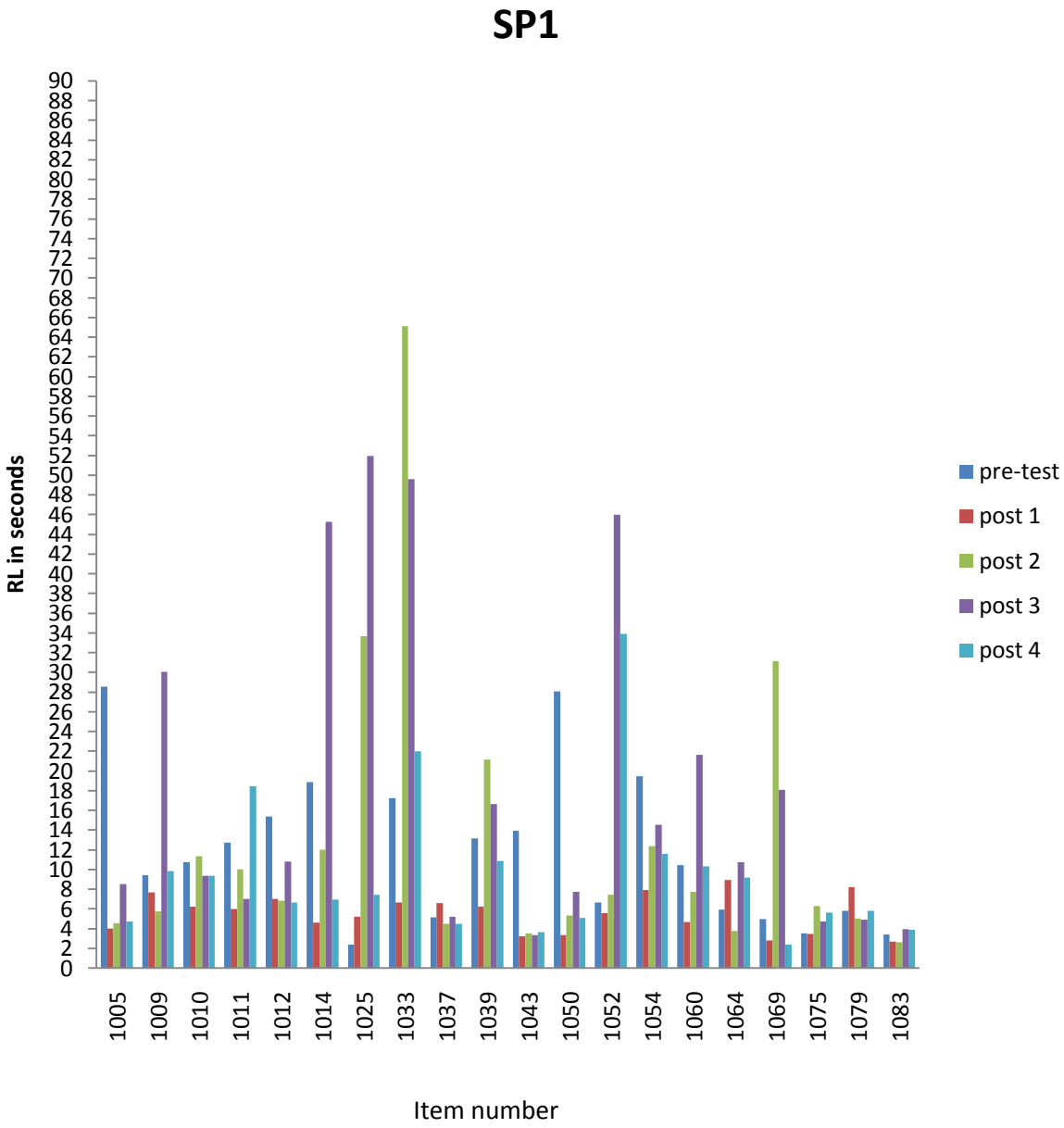


Figure 8.C.continued. Response latency for each item of the SP1 set in all tests for participant C.

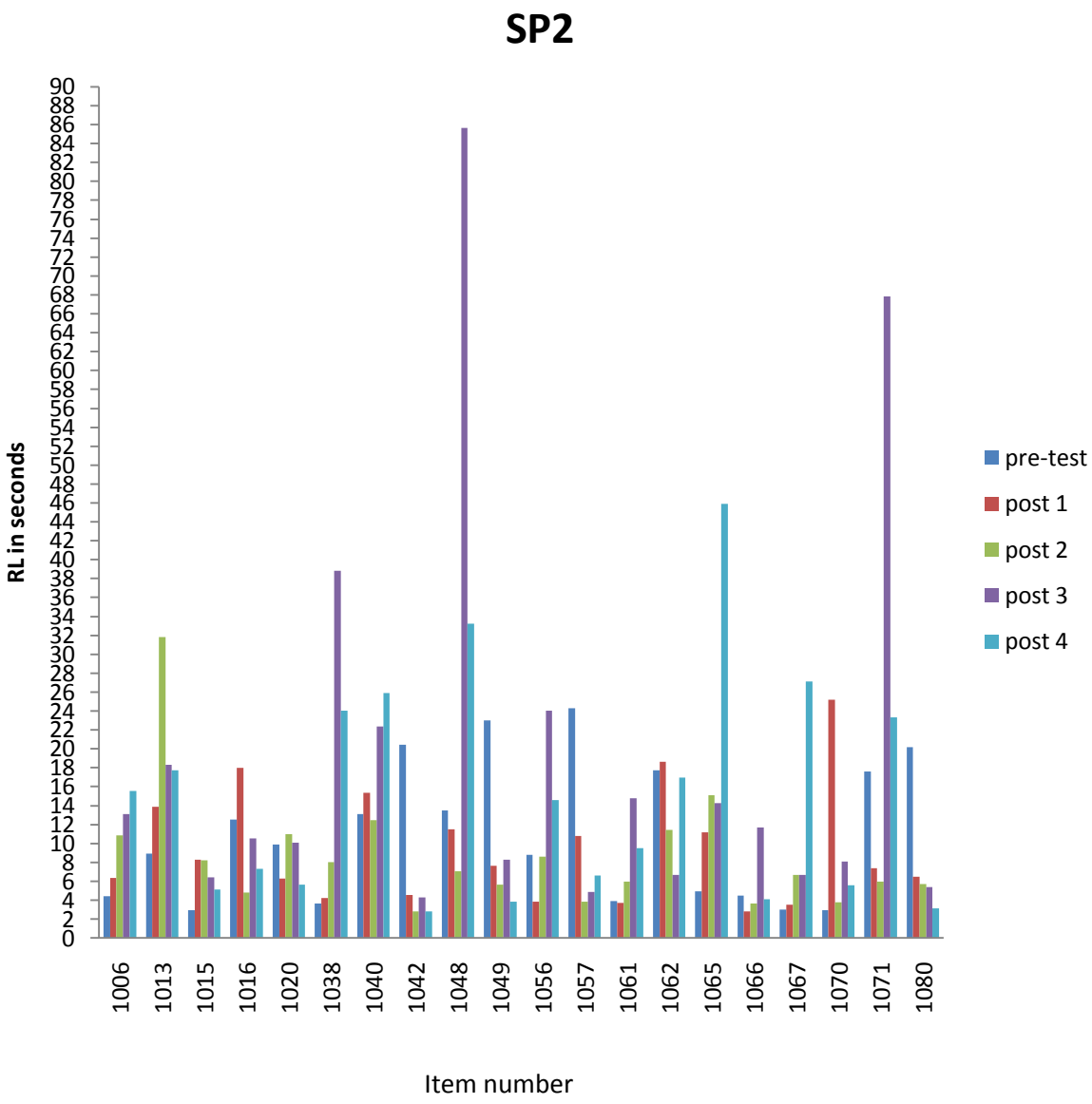


Figure 8.C.continued. Response latency for each item of the SP2 set in all tests for participant C.

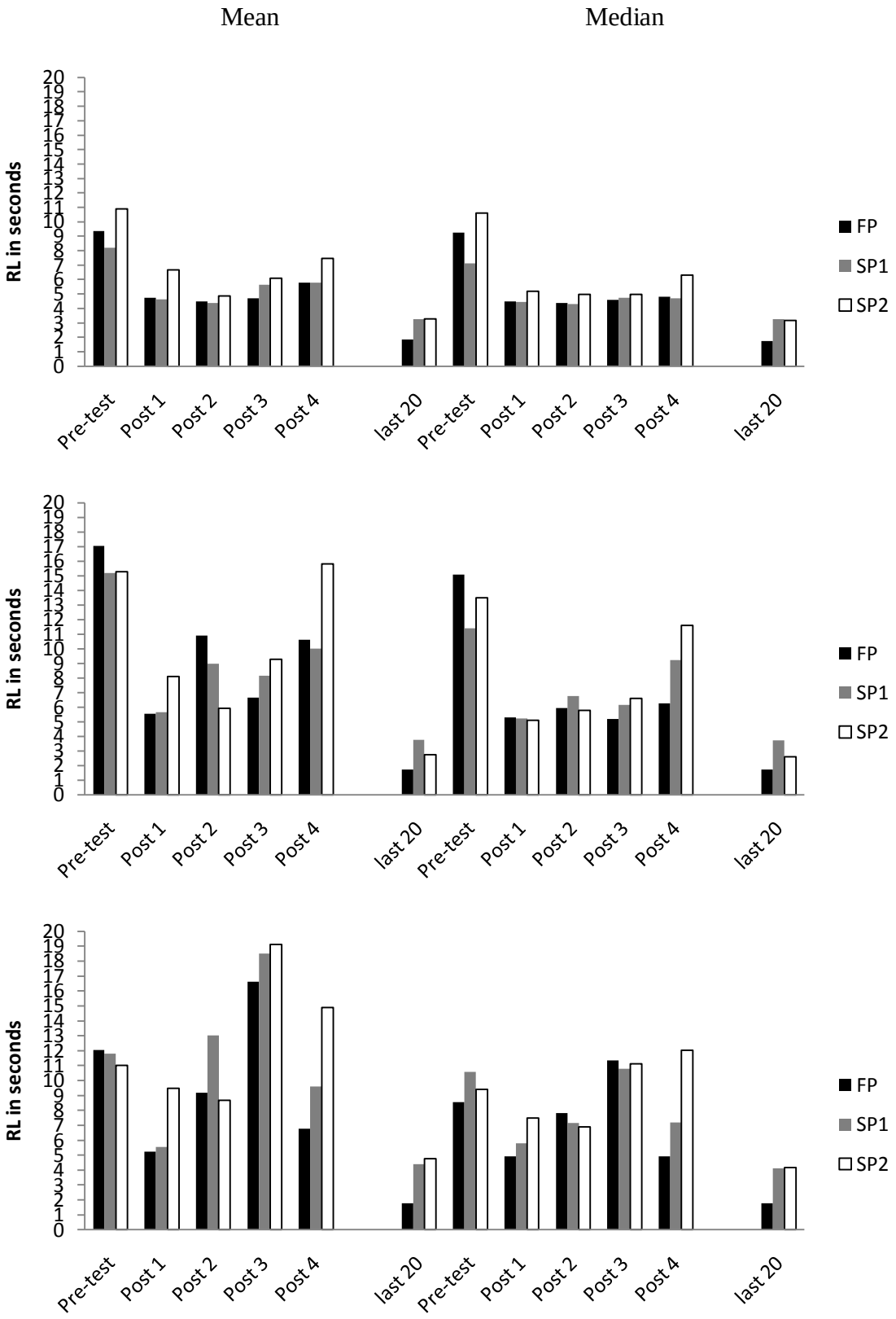


Figure 9. Mean and median response latencies (RL) for participants La (top), R (middle) and C (bottom) for each practice sets across tests (vs. mean and median RLs for the last 20 sessions).

RLs increased from post-test 1 to post-test 4 for La and R for all the 3 practice sets; but for participant C, they increased from post-test 1 to post-test 3 then decreased in post-test 4. The degree of change in mean and median RLs differed between practice sets and across participants. For example, much smaller change in mean and median RLs was observed between post-tests for FP and SP1 than SP2 for La. While there were smaller changes in median across tests in FP than in SP1 and SP2 for R, mean data are more variable with the means being affected by the extreme RLs seen in Figure 9. Participant C had similar degree of change across post-tests for all the practice sets.

Figure 10 presents the same data as Figure 9 but plotted across tests to show differences in mean and median RLs between practice sets. There was only a small difference in RLs between practice sets after all practices had been done (i.e., Post-test 2) for participant La; and SP2 had longer mean and median RLs than FP and SP1 only in post-test 4. Looking only at median RLs for R (Figure 11.B), the differences between practice sets remained small from post-test 1 to 3 but in post-test 4 the differences became larger because of the increased median RLs for SP1 and SP2. For C (Figure 11.C), there were only small differences between RLs for all practice sets in post-test 2 and 3 whereas there were large differences between median RLs for FP and SP2, SP1 and SP2 and a relatively small difference between FP and SP1. Overall, no big differences between practice sets was observed up until post-test 4, with FP having the shortest median RL, SP1 having the second and SP2 the longest.

In summary, accuracy was the same for all practice sets at the end of both practice phases but responses to items from SP1 and SP2 reached 100% accuracy in fewer trials than those from FP. At the end of the practice phases, response latencies were shorter in FP compared to SP1 and SP2 but they were similar for SP1 and SP2.

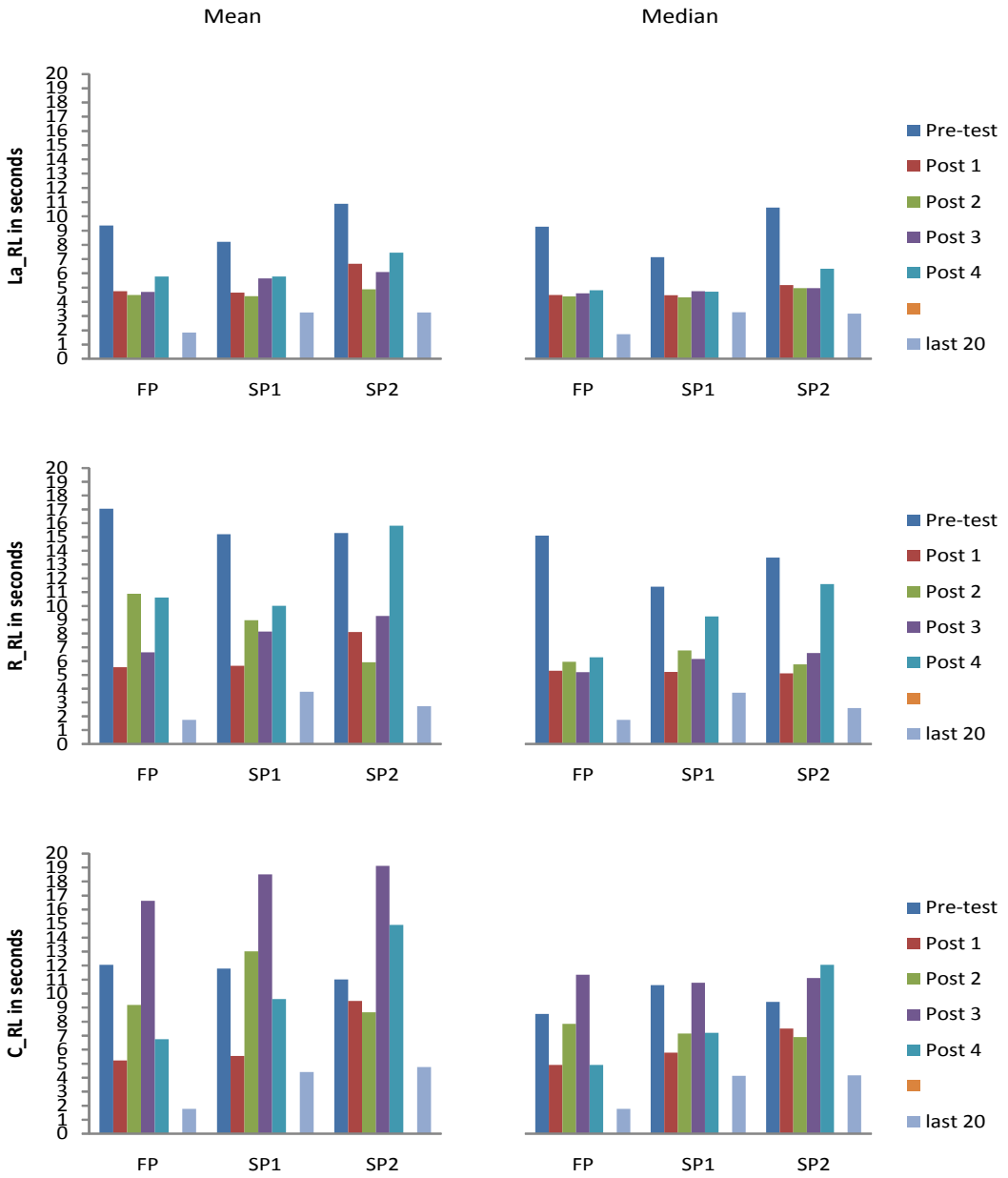


Figure 10. Mean and median RLs for the practice sets across tests for La (top), R (middle) and C (bottom).

Furthermore, accuracy was similar and high across all practice sets from post-test 2 to post-test 4. Response latencies increased over post-tests but did not differ over practice sets in post-tests 1 to 3. There was a great increase in response latencies for SP1 and SP2 in post-test 4, thus the difference in RL between FP and SP1 and between FP and SP2 became larger but the difference in RL between SP1 and SP2

remained small; and the SP2 set had the longest response latency amongst the 3 practice sets in post-test 4. Although items for the SP2 set had not been practiced in post-test 1, the median response latency for this set decreased to similar level as those or FP and SP1 while accuracy for the SP2 set remained low.

Discussion

One aim of the current study was to examine whether rate-building practice would produce higher accuracy than non rate-building practice and another was to examine whether undergoing the rate-building practice had any influence on the response latency to complete a response for the items that had been practiced only slowly. The results showed that, both rate-building and non rate-building practices, when equated for the amount of practice, produced the same high accuracy in remembering statistics definitions; and the high accuracy retained equally well after 4 weeks having no practices. Although response latencies differed between the two methods during the practice period, they did not differ in the tests where high rate was not required. Response latencies decreased from the beginning to the end of the practice periods for all practice sets, including for the two sets that were practiced slowly. The findings are discussed below.

Extensive practices, either in a fluency-building manner or in a slow manner, resulted both in high accuracy and in the information being practiced being retained after 4 weeks of no practices. The accuracy levels were similar, regardless of the type of practice, and this result is consistent with the previous within-subject studies that found no difference in accuracy between the two methods of practice immediate after practice and up to 8 weeks after practice (Wheetley, 2005; McGregor, 2006; Clark, 2007). This is also consistent with the between-subject studies (Fox & Ghezzi, 2003; Peladeau, Forget & Gagne, 2003) which found no differences between the groups that used different methods of learning. Yet, it is not consistent with the widely reported finding in precision teaching that compared to other methods of learning, rate-building practice produces longer lasting higher level of accuracy(e.g. Beneke, 1991; Le Grice, Mabin, & Graham, 1999; Bucklin, Dickenson, & Brethower, 2000; Johnson

& Layng, 1994; Singer-Dudek & Greer, 2005). These studies, as previously reported, did not control the amount of practices between rate-building and the comparison method. Therefore, the higher level of accuracy reported in these studies may simply be due to the larger amount of practice, as suggested by Doughty et al. (2004).

A greater number of trials was required to reach 100% accuracy in fast (rate-building) practice than for both slow (non rate-building) practices. This result is consistent with Wheetley (2005), McGregor (2006) and Clark (2007)'s studies. However, data from the Learning to Accuracy Phase showed that there was a slight tendency for the number of trials required to meet criterion to be larger for some of the items in the SP1 and SP2 sets than it was for the items in the FP set across participants. It is possible that having larger number of exposures to the SP1 and SP2 sets in the pre-training phase reduced the number of practice trials required to reach 100% accuracy. Although McGregor (2006) and Clark (2007) both equated the number of correct responses for all items before their extensive practice periods, larger numbers of trials were also required for the fast practice set to reach accuracy criteria than were required for the slow practice set during practice. McGregor (2006) suggested that the reasons for this could be because learning to fluency was a more difficult task or it was a new way of learning which the participants were not familiar with. It is not clear whether the same results would have been obtained if the number of practice trials had been equated for all items in the current study. One way of resolving this would be to equalise the number of correct responses in the Learning to Accuracy Phase. Another solution would be to divide the definitions into sets after Learning to Accuracy Phase and before Practice I Phase so that the definitions could be allocated into groups based on the total number of learning trials. Thus it might have been preferable to allocating them according to the time taken to read each item

as it was done in the current study. It should be noted that if this were to be used, the sets for each individual participant have to be different and the computer program would be harder to create.

The fast practice (FP) and the slow practice (SP1) that occurred together differed in the rate of presentation of trials but equated the number of practice trials and the number of feedback events. The feedback for fast practice was being told the rate of correct responses and the graphing of this on the semi-log charts; the feedback for slow practice was the number of correct responses and graphing accuracy (% correct). Both slow practices used the same rate of presentation and the same number of practice trials, but differed in the amount of feedback. That is, the second slow practice gave only the number of trials completed at the end of a session, and no feedback on total correct and no graph. However, the accuracy was high for all the three sets of definition immediate after practice and for up to 4 weeks after no practice. This may suggest that although the amount of feedback had been suggested to influence learning outcomes (Doughty et al., 2004), the types and amount of feedback had no effect on the retention of accuracy in the current study.

This study, as discussed earlier, recorded the response latency to complete a response for each item so that the actual response speed for the participants could be examined more closely. In Clark's (2007) thesis, the rate of responding was calculated by dividing the time taken for the slowly practiced items by the number of items completed. This did not reflect how quickly participants responded to the items from the slowly practiced set since there was the imposed time frame of 6 s for each response for the slow practice condition. The response latencies obtained in the current study referred to the duration between the time a term was presented and the time when the space bar was hit as an indication of having had answered that item;

these response latencies did not reflect the rate of responses for SP1 and SP2 because the presentation rates were controlled by the computer.

The procedure is argued to be an improvement on the designs used by Wheetley (2005), McGregor (2006) and Clark (2007)'s studies in that current study had succeeded in ensuring participants took longer to completed responses for the items in the slow practice sets than for the items in the fast practice set. Response latencies were shorter for the fast practice than for the two slow practices; and they were not different between the two slow practices. The median response latency for the slow practices (close to 4 s) was much longer than the median response latencies required to achieve the rate aim in the fast practice (up to 2 s). The differences are quite large but did not result in different learning outcomes in the levels of accuracy.

When examining the data from within the practice phases only (Practice I Phase and Practice II Phase), it was found that response latencies decreased with the increased number of practice trials for the two sets for fast and slow practices occurring together (FP and SP1) as well as for the set for slow practice that occurred alone (SP2). Interpreting the similar median response latencies for the two slow practices observed in the last 20 sessions of the practice phases (Practice I Phase and Practice II Phase) has to be done with caution because it is not known whether or not they had been influenced by the fast practice that occurred alongside the first slow practice (SP1). Since the median response latencies for the last 20 sessions for SP1 was the same as those for SP2, it could be either there was no carryover effect from FP or that both of them were affected by FP equally. If there was no carryover, then the fact that response latencies in both of the slow practices decreased (even though this was not required by the procedure) suggests that practice beyond accuracy alone without a speed requirement made the responses more fluent.

However, a carryover from fast practice cannot be ruled out because Practice II Phase (slow practice 2) occurred very close in time to Practice I Phase (FP and SP1) and over a very short period (with a large amount of practices in any one session). This was done so that the later retention of accuracy for the three sets of definitions could be assessed at the same time. However, this meant that the time between these two phases (Practice I and II Phases) may have been too short to minimise the influence of the FP on SP2. That is to say, response latencies for SP2, although it occurred alone, might have decreased because of the influence of the earlier FP. Therefore, whether or not the response latencies for the slow practices were affected by the fast practice cannot be determined. If there was carryover, the shorter response latencies observed at the end of the practice phases for SP1 and SP2 cannot be attributed solely to the effects of practice; it could be the result of either the influence of FP or the amount of practice, or both of these two factors. Scheduling the practice phase with the SP2 set to take place after a longer delay following FP and SP1, for example, after 4 weeks, and testing the retention for accuracy for this set separately may help eliminating any carryover effect.

The median response latencies increased from the last 20 sessions of training to post-test 1 for all 3 sets of definitions. The increase was expected for the FP set because participants were no longer required to response quickly in post-test 1. But the increase for SP1 and SP2 was unexpected since they were instructed to response slowly in the training and no speed was required in all the tests. This means that some other factors may have influenced the decrease in response latencies for the last 20 sessions for the two slow practice sets. It could be the training of fast practice as discussed earlier; or it could be resulted from the design of the procedure which imposed a 5 s limit for the display of definition in the practice phases. Specifically,

there was a 5 s time limit for reviewing the definition to each term in the practices. The 5 s limit was hoped to be long enough for participants to finish reading all words in one definition out loud but not enough for reading it for more than once to prevent participants making more practice than those were recorded. However, at the beginning of the practice phases, participants could not finish reading all words in the definition within 5 s for some items on their first appearance. Participants then sped up to try finish reading the in a number of subsequent trials for those items.

Furthermore, because practice of the set of definition for the slow practice that occurred alone (SP2) was, it was hoped, going to be free of carryover effect from fast practice, the response latencies for this set were expected to be the same in pre-test and post-test 1. However, response latencies decreased greatly from pre-test to post-test 1 for all sets of definitions to a similar level. There are at least two possible explanations for this. The first possibility is that fast practice influenced the speed of responding for all three sets of definition not only during the practice period but also in the post-tests. Thus, learning to be fast generalized to SP2 set but note although fast, they were not accurate. It has been shown in the behavioural momentum literature on academic learning that when high preference (High P) behaviours are added before low preference (Low P) behaviours, responses to Low P behaviours will increase (Neef, Iwata, & Page, 1977; Cook, Guzaukas, Pressley, & Kerr, 1993; Cooke & Riechard, 1996; Belfiore, Lee, Vargas & Skinner, 1997; Cates & Skinner, 2000). Particularly, Cooke, Guzaukas, Pressley, and Kerr (1993) found that students had a greater rate of learning (calculated by dividing the number of items by the time taken to complete the items) for the unknown mathematics facts if they were interspersed with known facts in a 30/70% (unknown/known) ratio. Completing more items in a fixed time would mean participants responded faster to the items overall. Belfiore et

al. (1997) taught adolescents who had compliance and homework completion problems to learn mathematics problems in a procedure that interspersed 3 high preference single-digit (easy) problems with 1 low preference multi-digit (difficult) problems. They found that participants responded faster to the Low P problems when they were interspersed with the High P problems, compared to when the Low P problems were tested alone in the baseline condition. Thus, it is possible that such an effect was also present here. Two thirds of the items in post-test 1 had been extensively practiced and they were mixed with the one third that had not been practiced, thus answering the ones that had been practiced in post-test 1 may have built up the momentum in responding and this could be resulted in the decrease in median response latencies for the unpractised items. This is supported by the fact that although participants responded as quickly to the unpractised definitions, the responses for these were not as accurate as for the other two sets. There are at least two steps to examine this further. First of all, as discussed earlier, SP2 could be scheduled after a longer delay to eliminate the possibility of carryover effect from FP to SP2. Secondly, the 3 sets of definitions could be tested separately but in the same manner as here (un-timed and no rate requirement) in order to eliminate the possibility of fast responding carrying over within the tests.

Different from Wheetley (2005), McGregor (2006) and Clark (2007), the current study used only a slow and un-timed testing procedure in order to minimise the effect of fast-practice on the slowly-practiced items in regard to response latency. The response latencies were the same or very similar across practice sets (although different across participants) in pre-test, post-test 1 to 3 and this suggests that when not prompted to response quickly, participants responded at the same rate for all items regardless of the method with which they had been practiced. In post-test 4, 4 weeks

after training, the time it took the participants to respond stayed the same for the fast practice set while it increased greatly for the two sets practiced slowly. Thus, this suggests that although there was no difference in accuracy across the practice sets, participants maintained the fast response to the FP set better than they did to the SP1 and SP2 sets. The question of whether the accuracy and response latencies would differ if the delay to testing had been longer arose; more delayed post-tests (e.g., 8 and 12 weeks) in the future studies may help clarify this.

Limitations

Current study intended to adopt some aspects of previous studies and improve on them to examine whether accuracy and retention differed between sets of definitions practiced under the rate-building method and the non rate-building method. It also explored the possibility of a carryover effect from rate-building practice to the non rate-building practice. A number of limitations prevented a firm conclusion from the current study. In spite of the three previously mentioned limitations, which are having unequal number of learning trials before the practice period, scheduling SP2 to occur too close in time to FP and SP2 and the possibility of the influence of behavioural momentum, there are two more areas that need improvement.

Firstly, there was a high withdrawal rate in the current study, 4 out of the 7 participants who started the study did not finish either because it took them too long or because they did not want to come to university after term finished. The main reason that had increased the time required was because of their low rate of attendance to the practice sessions. The low attendance resulted in these participants having less frequent practices during a week, which lengthened the time needed for them to reach the accuracy criterion in the Learning to Accuracy Phase as well as

achieving the rate-aim in the Practice I Phase. Because incorporating rate-building practice required extra effort from participants, it might help increasing commitment to the study and attendance by making it clear there would be reward for these extra efforts. It could be done by making a clearer statement on the information sheets that were handed out to the participants or by some other form of rewards.

Secondly, some of the items, predominantly in the set for fast practice, had shorter response latencies in the first 20 sessions than in the last 20 sessions in the practice phases. These extremely short latencies kept increasing after the first 20 sessions before decreasing again. After a closer look at the data, it was found that this resulted from participants skipping these unfamiliar items as a technique of getting more items in one minute. These responses were incorrect and thus the response latencies did not truly reflect the latencies in which participants were able to respond. It would be better if the computer program added a function that matches each response with its accuracy and response latency to allow for examination of response latencies separately for correct and incorrect responses.

Current study aimed at comparing the effects equal amounts of slow and fast practice on the maintenance of accuracy after the end of practice. Both of fast and slow practice produced high accuracy and the accuracy remained high for both after 4 weeks. This result adds support to the literature that suggests extended practice without rate building is sufficient in producing high accuracy---contrary to the precision literature that claims rate building is necessary for the superior learning outcomes of Precision Teaching in terms of accuracy. The current study succeeded in producing different speeds of practice for the different sets of definitions and so, in comparison to the the three unpublished theses mentioned in the introduction, gives stronger support to the result on accuracy. However, the possibility of a carryover

effect from fast practice to slow practice could not be eliminated. This, along with the other limitations described earlier prevents a firm conclusion from current study. Further refinements of the research design are needed to examine or eliminate any carryover effects. The possibility of generalization of fast practice from the practice to the tests may also be examined further.

References

AimChart (2008). *Free PT Charting on the Web.* Retrieved December 1th, 2009, from http://www.aimchart.com/index.shtml.

Belfiore, P.J., Lee, D.L., Vargas, A.U., & Skinner, C.H. (1997). Effects of High-Preference Single-Digit Mathematics Problems Completion on Multiple-Digit Mathematics Problem Performance. Journal of Applied Behavior Analysis, 30, 327-330.

Beneke, W.M. (1991). Precision Teaching to Enhance Reading Skills of Introductory Psychology Students. *Journal of Precision Teaching, 8,* 37-43.

Binder, C. (2004). A Refocus on Response-Rate Measurement: Comment on Doughty, Chase, and O'Shields. *The Behavior Analyst, 27,* 281-286.

Binder, C., Haughton, E., and Bateman, B. (2002). *Fluency: Achieving True Mastery in the Learning Process*. Retrieved March 28, 2008, from http://curry.edschool.virginia.edu/go/specialed/papers/binder-et-al_fluency.pdf.

Bucklin, B.R., Dickenson, A.M. & Brethower, D.M. (2000). A comparison of the effects of fluency training and accuracy training on application and retention. *Performance Improvement Quarterly, 13,* 140-163.

Cakes, G.L., & Skinner, C.H.(2000). Getting Remedial Mathematics Students to Choose Homework with 20% and 40% More Problems: An Investigation of the Strength of the Interspersing Procedure. Psychology in the Schools, 37, 339-347.

Chapman, S.S., Ewing, C.B. & Mozzoni, M.P. (2005). Precision Teaching and Fluency Training across Cognitive, Physical, and Academic Tasks in Children with Traumatic Brain Injury: A Multiple Baseline Study. *Behavioral Intervention,* 20, 37-49.

Clark, T. (2007). The Relative Effects of Fluency and Accuracy-Based Instruction on Learning Performance Outcomes. Unpublished master's thesis, University of Waikato, Hamilton, New Zealand.

Cooke, N.L., Guzaukas, R., Pressley, J.S., & Kerr, K. (1993). Effects of Using Ratio of New Items to Review Items during Drill and Practice: Three Experiments. Education and Treatment of Children, 16, 213-234.

Cooke, NL., & Reichard, S.M. (1996). The Effects of Different Interspersal Drill Ratios on Acquisition and Generalization of Multiplication and Division Facts. *Education and Treatment of Children,* 19, 124-142.

Cooper, J.O., Heron, T.E., and Heward, W.L. (2007). *Applied Behavior Analysis.* Upper Saddle River, J.J.: Pearson/Merril-Prentice Hall.

Darvel, A., & Rose, D. (2006). Fast and Fluent or Careful and Correct? The relative Effects of Fluent and Accurate Practice on Acquisition, Maintenance and Generalization. Unpublished Master's thesis, University of Auckland, Auckland.

Doughty, S.S., Chase, P.N., and O'Shields, E.M. (2004). Effects of Rate Building on Fluent Performance: A Review and Commentary. The *Behavior Analyst, 27,*7-23.

Downer, A. (2007). The National Literacy Strategy Sight Recognition Programme Implemented by Teaching Assistants: A Precision Approach. *Educational Psychology in Practice,* 23, 129-143.

Eastridge, D., and Mozzoni, M.P. (2005). Fluency and Functional Motor Skills Following Brain Injury. *Behavioral Interventions, 20,* 77-89.

Evans, S.S. and Evans, W.H. (1985). Frequencies that Ensure Skill Competency. *Journal of Precision Teaching,* 1, 25-30.

Fox, E.J. and Ghezzi, P.M. (2003). Effects of Computer-based Fluency Training on Concept Formation. *Journal of Behaivoral Education, 12,* 1-21.

Hartnedy, S.L., Mozzoni, M.P., and Fahoum, Y. (2005). The Effect of Fluency Training On Math and Reading Skills in Neuropsychiatric Diagnosis Children: A Multiple Baseline Design. *Behavioral Interventions, 20,* 27-36.

Howell, D.C. (2008). *Fundamental Statistics for the Behavioral Sciences.* Belmont, CA : Thomson Wadsworth.

Howell, K.W., & Lorson-Howell, K.A. (1995). What's the Hurry? Fluency in the Classroom. *Journal of Precision Teaching, 7, 24-28.*

Johnsoon, K. R., and Layng, T.V. (2004). The Morningside Model of Generative Instruction. *Behavior analysis in education: Focus on measurably superior instruction.* pp. 173-197.

Le Grice, B., Mabin, T., & Graham, S. (1999). I Can Do Maths: Changing Children's Mathematics Percentile Rankings. In *Proceedings from the Combined Meeting of the Australian Association for Research in Education and the New Zealand Association for Research in Education.* Melbourne, Australia.

Lindsley, O.R.(1991). Precision Teaching's Unique Legacy from B.F. Skinner. *Journal of Behavioral Education, 1,* 253-266.

Merbitz, C., Vieitez, D., Merbitz, N. H., and Binder, C. (2004). Precision Teaching: Applications in Education and Beyond. In D.J. Moran & R.W Malott (Eds), *Evidence-Based Educational Methods* (pp. 63-78).

McGregor, S.J. (2006). Practice Makes the Difference: The Effect of Rate-Building and Rate-Controlled Practice on Retention. Unpublished master's thesis, University of Waikato, Hamilton, New Zealand.

Peladeau, N., Forget, J., and Gagne, F. (2003). Effects of Paced and unpaced Practice on Skill Application and Retention: How Much Is Enough? *American Educational Research Journal, 40,* 769-801.

Pocock, T.L. (2006). An Analysis of Precision Teaching. Unpublished Ph.D. thesis, University of Waikato, Hamilton, New Zealand.

Singer-Dudek, J. and Greer, R.D. (2005). A Long-Term Analysis Of the Relationship Between Fluency and the Training and Maintenance of Complex Math Skills. *The Psychological Record, 55,* 361-376.

Walker, V.L. (2008). Effects of rate of feedback and reinforcement on novel responding. Unpublished doctoral thesis, West Virginia University, Morgantown, West Virginia, United States.

Appendix 1

Participant Information Sheet

Title of project: Fast practice or merely more practice results in better learning?

Researcher: Kong Xiuyan (KITT) (kitt_kong@hotmail.com)

Supervisor: Prof. Mary Foster (m.foster@waikato.ac.nz)

My name is Kong xiuyan (Kitt), and I am a student at the University of Waikato. I am studying to complete my masters in psychology. Prof.Mary Foster is my supervisor for this project. This information sheet is to let you know about this research and invite you to participate.

We want some university students to learn information related to a subject they are studying and then for them to use two different methods of practice with different parts of the material they have learned to compare the two methods. If you take part in this study, you will first be asked to work with me to choose a subject area that you would like to improve on and then I will consult with the teachers to identify key concepts that will be useful to you in both the short term and the long term. The concepts will be taught to you first and then you will practice in either a fast way or slowly using a computer to present the concepts – while you respond out loud - with me present.

During the study we will arrange sessions of about 30 minutes long for as many of the five week days each week as you can make - for up to five weeks – in which you will practice the terms using a computer. Your performance will be recorded on a video camera without showing your face. The video recording will be used only for checking the data's reliability and for no other purpose. You will be tested some time during the study on how well you know the terms. There will be an additional three sessions the times of which will be arranged with you either in person or through email at 2, 4 and 8 weeks after the practice period has ended to allow us to assess the effectiveness of the learning. You will also be asked to fill out a brief questionnaire about your experience during the study. This research will occupy approximately 2.5 hours of your time per week.

Information about your performance will be kept so it is anonymous and so there will no way you can be identified as a participant in any reports of this study. You can choose whether or not you want to take part in this study, and do not have to give a reason if you choose not to take part. If you change your mind later and decide to pull out of the study, you can withdraw without consequences at any time. All of the information I collect during the study will be kept in a locked cabinet in the psychology department, only my supervisor and I will be able to access this information.

If you have any further questions please feel free to ask them.

Appendix 2

University of Waikato

Psychology Department

CONSENT FORM

PARTICIPANT'S COPY

Research Project: **Fast practice or merely more practices?**

Name of Researcher: Xiuyan Kong (kitt)

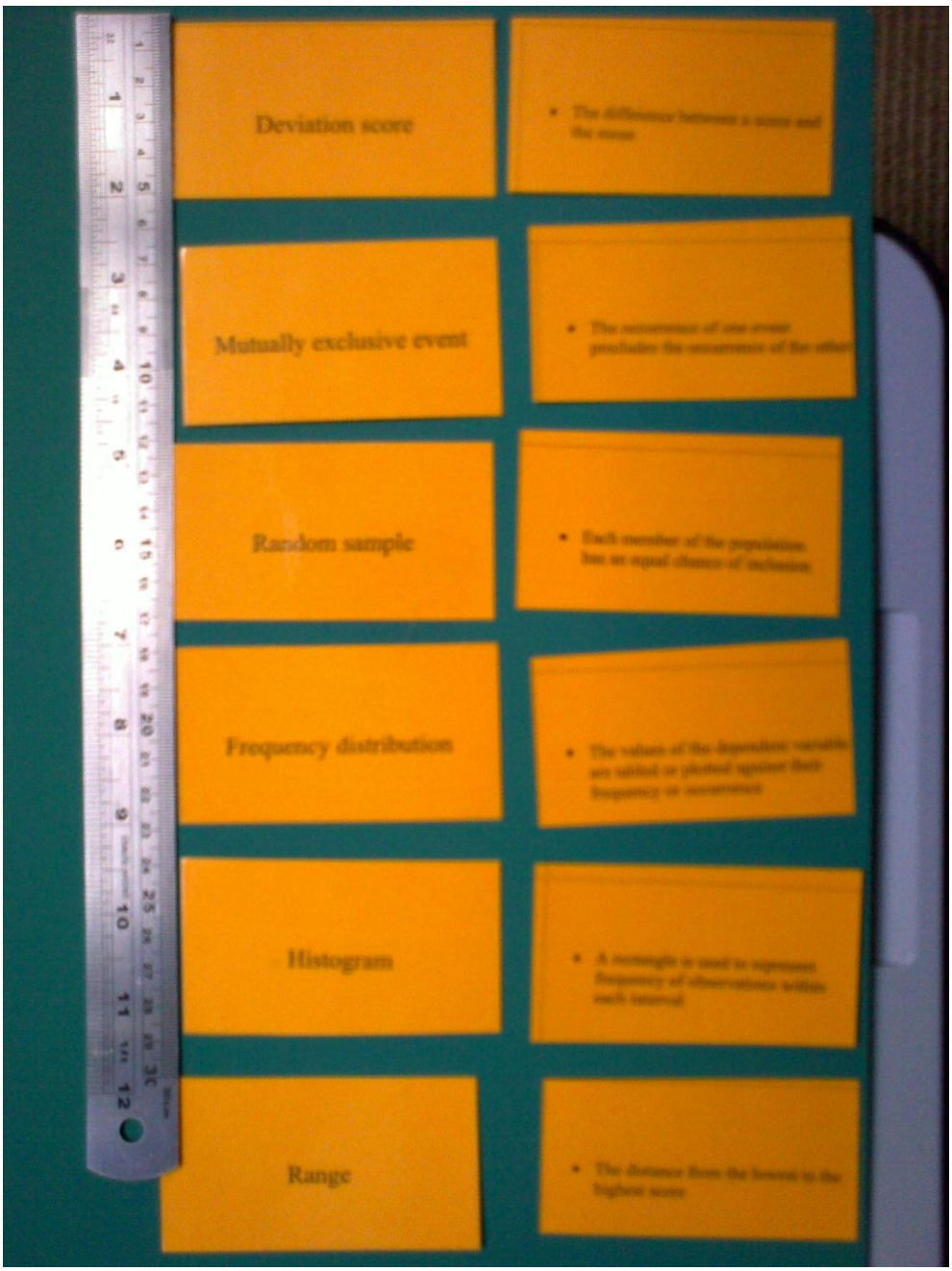
Name of Supervisor (if applicable): Mary Foster

I have received an information sheet about this research project or the researcher has explained the study to me. I am aware that my performance during the sessions will be video-taped without showing my face. I have had the chance to ask any questions and discuss my participation with other people. Any questions have been answered to my satisfaction.

I agree to participate in this research project and I understand that I may withdraw at any time. If I have any concerns about this project, I may contact the convenor of the Research and Ethics Committee (Dr Robert Isler, phone: 838 4466 ext. 8401, e-mail r.isler@waikato.ac.nz)

Participant's
Name:_____Signature:_____Date:_____

Appendix 3


Deviation score
Mutually exclusive event
Random sample
Frequency distribution
Histogram
Range

ID	SET	TERM	DEFINITION
1001	1	Population	Complete set of events of interest
1002	1	Correlation coefficient rho (p)	The correlation coefficient for the population
1004	1	Statistics	Characteristics of a sample
1018	1	Variables	Properties of objects or events that can take on different values
1023	1	Continuous variables	Take on any values
1028	1	Deviation score	The difference between a score and the mean
1032	1	Random assignment	The allocation or assignment of participants to groups by a random process
1041	1	Symmetric distribution	Having the same shape on both sides of the centre
1053	1	Mode	The most commonly occurring score
1055	1	Mean	The sum of the scores divided by the number of scores
1058	1	Sample	A set of observations; subset of a population
1059	1	Exhaustive events	A set of events that represents all possible outcomes
1063	1	Sample variance	Sum of squared deviations about the mean divided by N-1
1068	1	Standard normal distribution	A normal distribution with a mean equals to 0 and a standard deviation equal to 1; denoted as N(0,1)
1072	1	Two-tailed test	Rejects extreme outcomes in either tail of a distribution
1076	1	One-tailed test	Rejects extreme outcomes in one specified tail of a distribution
1077	1	Sampling with replacement	The item drawn on trial N is replaced before the next draw
1078	1	Parameter	Measure of a population characteristic; Usually estimated, rarely computed
1081	1	Mutually exclusive event	The occurrence of one event precludes the occurrence of the other
1082	1	Type I error	Rejecting null hypothesis when it is true

ID	SET	TERM	DEFINITION
1005	2	Random sample	Each member of the population has an equal chance of inclusion
1009	2	Spearman's correlation coefficient (r)	For rank-ordered data
1010	2	Categorical data	Data representing counts or number of observations in each category
1011	2	Quantitative data	The assignment of numbers to objects
1012	2	Scales of measurement	Characteristics of relations among numbers assigned to objects
1014	2	Ordinal scale	Numbers used only to place objects in order
1025	2	Sampling error	Variability of a statistic from sample to sample due to chance
1033	2	Frequency distribution	The values of the dependent variable are tabled or plotted against their frequency of occurrence
1037	2	Expected values	The long-range average of a statisticaly repeated sample
1039	2	Line graph	The Y values corresponding to different values of X are connected by a line
1043	2	Unimodal distribution	Having one distinct peak
1050	2	Positively skewed distribution	Trails off to the right
1052	2	Percentile	The point below which a specified percentage of the observations fall
1054	2	Median	The score corresponding to the point having 50% of the observations below it when the observations are arranged in numerical order
1060	2	Power	The probability of correctly rejecting a false null hypothesis
1064	2	Standard deviation	A measure of scatter; The deviation from the mean
1069	2	Directional test	A one-tailed test
1075	2	Standard error	The standard deviation of a sampling distribution
1079	2	Correlation coefficient	A measure of the relationship between variables
1083	2	Correlation	Relationship between variables

ID	SET	TERM	DEFINITION
1006	3	Decision tree	Graphical representation of decisions involved in the choice of statistical procedures
1013	3	Nominal scale	Numbers used only to distinguish among objects
1015	3	Interval scale	Equal intervals between objects represent equal differences
1016	3	Ratio scale	A scale with a true zero point; ratios are meaningful
1020	3	Discrete variables	Take on a small set of possible values
1038	3	Histogram	A rectangle is used to represent frequencies of observations within each interval
1040	3	Bar graph	The frequency of occurrence of different values of X is represented by the height of a bar
1042	3	Bimodal distribution	Having two distinct peaks
1048	3	Ranked data	Observations have been replaced by their numerical ranks from lowest to highest
1049	3	Negatively skewed distribution	Trails off to the left
1056	3	Conditional probability	The probability of one event will occur given the occurrence of some other event
1057	3	Range	The distance from the lowest to the highest score
1061	3	Quartiles	The points which break the distribution into fourths
1062	3	Variance	Measures scatter; Average of squared deviation from mean
1065	3	Bias	When a statistic is systematically different from the parameter it estimates
1066	3	Nondirectional test	A two-tailed test
1067	3	Boxplot	Graphical representation of the dispersion of a sample
1070	3	Z score	Number of standard deviations above or below the mean
1071	3	Standard scores	Scores with a predetermined mean and standard deviation
1080	3	Type II error	Not rejecting null hypothesis when it is false

Set 3-SP2

Appendix 5

Participant:

Date:

	Read	Correct (I)/Incorrect (--)
Discrete variables		
Histogram		
Interval scale		
Nondirectional test		
Nominal scale		
Negatively skewed distribution		
Quartiles		
Range		
Deviation score		
Ranked data		
Ratio scale		
Expected values		
Standard scores		
Variance		
Frequency distribution		
Line graph		
Median		
Exhaustive events		
Mutually exclusive event		
Mean		
Mode		
One-tailed test		
Population		
Parameter		
Random assignment		
Sample		
Standard normal distribution		
Two-tailed test		
Type I error		

Statistics		
Categorical data		
Correlation		
Correlation coefficient		
Symmetric distribution		
Standard deviation		
Unimodal distribution		
Correlation coefficient rho (p)		
Continuous variables		
Ordinal scale		
Percentile		
Positively skewed distribution		
Power		
Scales of measurement		
Quantitative data		
Sampling error		
Random Sample		
Standard error		
Spearman's correlation (r)		
Sample variance		
Variables		
Bimodal distribution		
Bar graph		
Z score		
Bias		
Boxplot		
Conditional probability		
Type II error		
Decision tree		
Directional test		
Sampling with replacement		

Appendix 6 (pp. 87- 90)

Power point illustration of the computer display for the Practice I & II Phases.

Welcome !

You will be practicing two sets of definitions previously learned, one as fast as you can and the other slowly and accurately.

There will be three fast practices followed by a brief break. Then the slow practice will start.

Press space bar for more instructions

1

Please say out loud the definition of the term presented on the screen and press the space bar once you have finished.

Press space bar for more instructions

2

The definition of that term will appear on the screen once you have pressed the space bar.

Press the space bar again when you have finished reviewing it and the next term will come up.

Press space bar for more instructions

3

We will start with the fast practice.

Say out loud definitions as fast and accurately as you can.

Each practice will last for 1 minute.

Press space bar for more instructions

4

Questions? Please ask ☺

Press the space bar to start when you are ready.

Start

5

behaviorism

Press space bar for Answer

6

behaviorism

The philosophy of behavior analysis

Press space bar when you have finished reviewing

7

Next

8

Times up!

Well done ☺

Take a short break and get ready for the next 1 minute fast practice.

9

We will start another 1 minute fast practice

Remember to work as fast and accurately as you can and try to beat your record from the previous practice.

Press the space bar when you are ready to start.

start

10

Times up!

End of fast practice. GOOD WORK ☺

You can have a 5 min break from now.

Relax, chat, laugh, or walk around.

The researcher will inform you when break-time is finished and instruction will appear on screen again.

11

12

We will proceed to slow practice now.

Please say the definitions out loud slowly, and as accurately as you can.

Computers will control the pace.

Press space bar for more instructions

13

Say the definitions out loud, then press the space bar and the answer will appear.

Press the space bar when you have finished reviewing it and wait for the presentation of a new term.

Carry on following the instructions.

Press space bar for more instructions

14

Questions? Please ask ☺

Press the space bar to start when you are ready.

Try to be as accurately as you can.

Start

15

Fluency

Press space bar for Answer

16

Fluency

Retention
Endurance
Stability
Application
Adduction
Problem solving

Press space bar when you have finished reviewing it

17

18

Good work! ☺

Take a short break and get ready for more slow practice.

19

20

We will start another slow practice.

Remember to work as accurately as you can and try to keep the good performance up.

Press the space bar when you are ready to start.

Start

21

Slow practice completed.

GOOD WORK ☺

End of session, thanks for your hard work. please go and see the researcher.

22

Appendix 7

Power point illustration of the computer display for the tests.

Welcome!

Press space bar for more instructions

1

Please say out loud the definition of the term presented on the screen and press the space bar once you have finished to bring up another term.

Press space bar to go to next page

2

Questions? Please ask ☺

Press the space bar to start when you are ready.

Start

3

behaviorism

Press space bar for next term

4

Applied behavior analysis

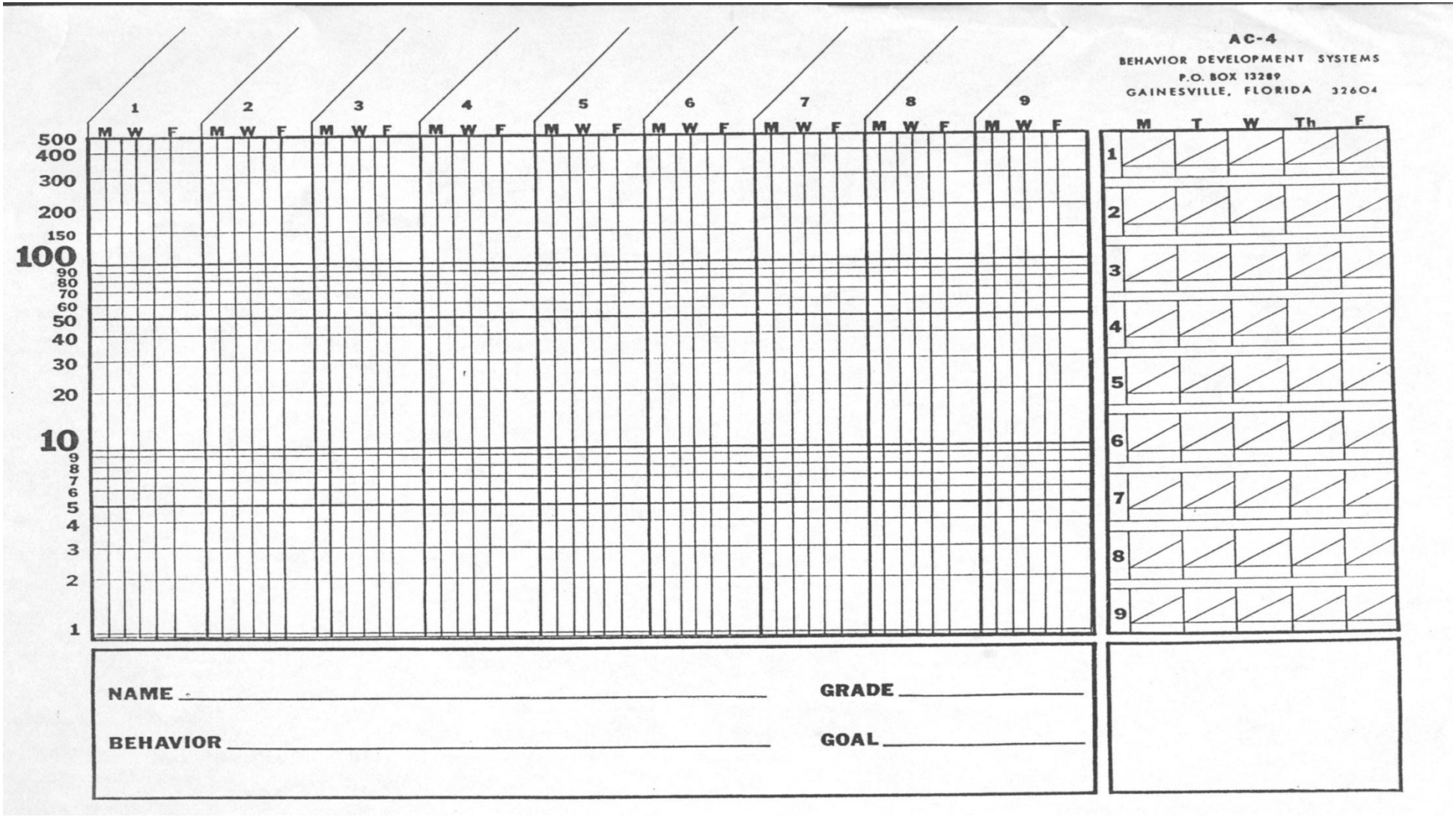
Press space bar for next term

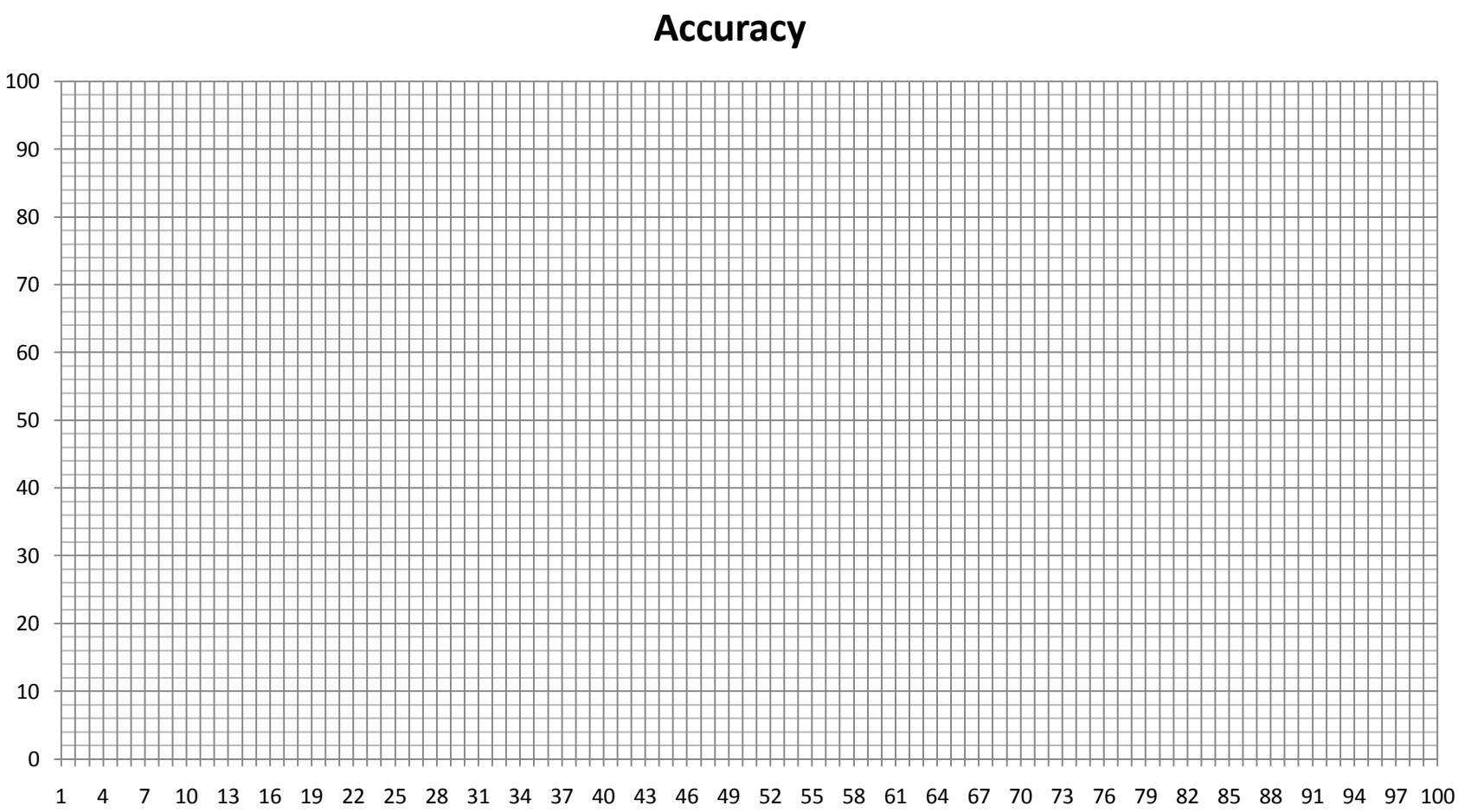
5

Well Done! ☺

End of session, thanks for your hard work.

6


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500
400
300
200
150
100
90
80
70
60
50
40
30
20
10
9
8
7
6
5
4
3
2
1
M T W Th F
1
2
3
4
5
6
7
8
9
NAME
GRADE
BEHAVIOR
GOAL


Accuracy
100
90
80
70
60
50
40
30
20
10
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 64 67 70 73 76 79 82 85 88 91 94 97 100

Appendix 10 (pp. 94-104)

The actual semi-log charts and percent correct charts used by the participants, with data.

3071- Participant La (A)

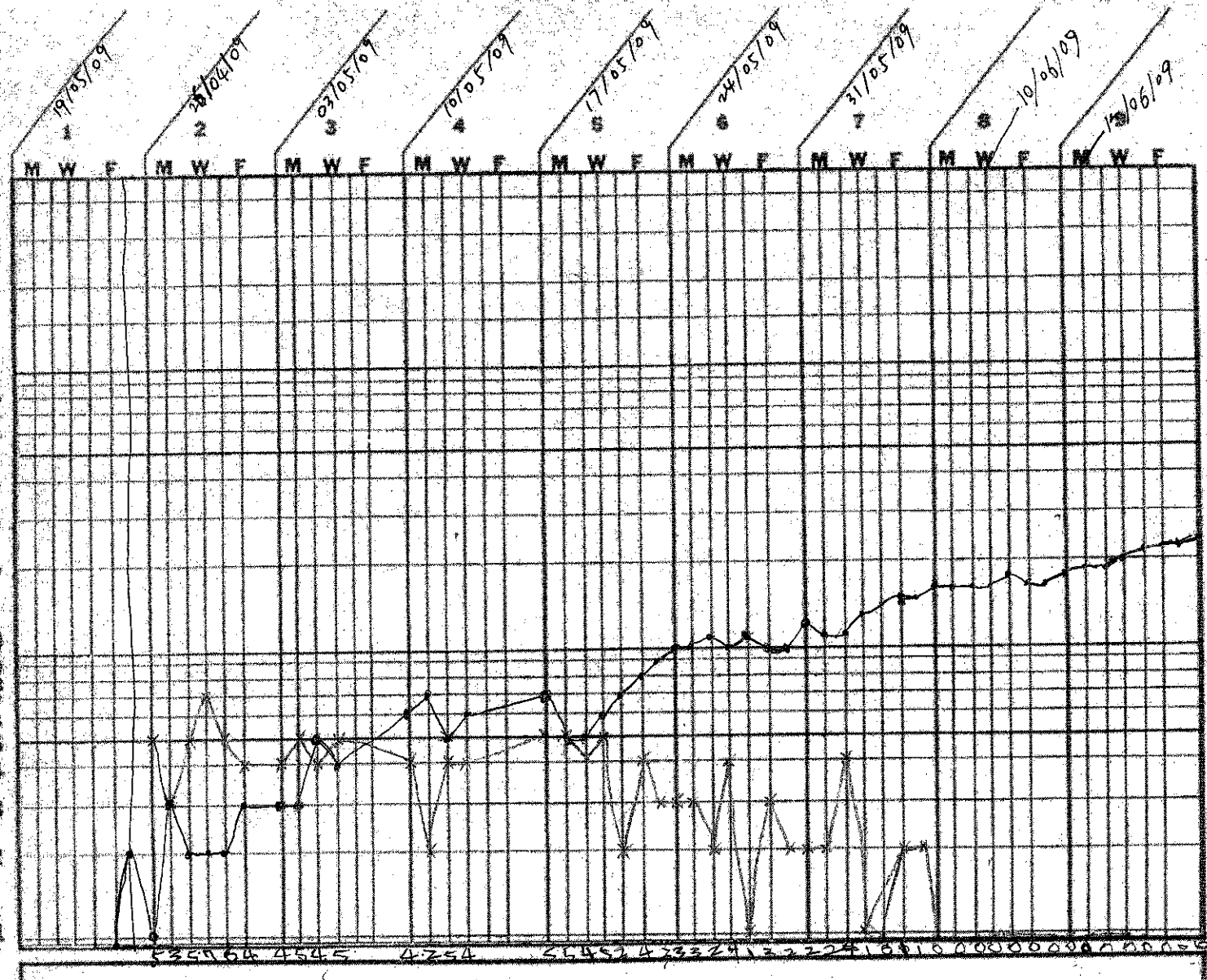
3074- Participant R (B)

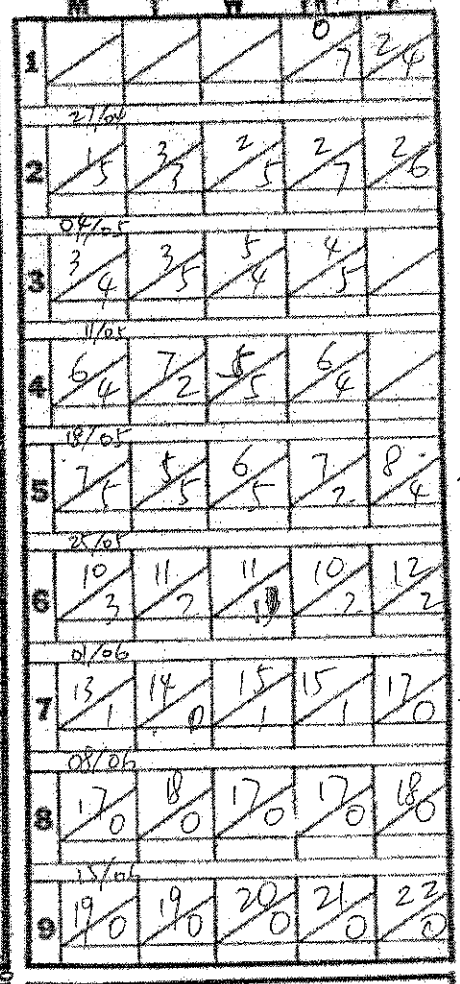
3065- Participant C (C)

3075- Participant T (D)

3069- Participant A (E)

AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13289
GAINESVILLE, FLORIDA 32604



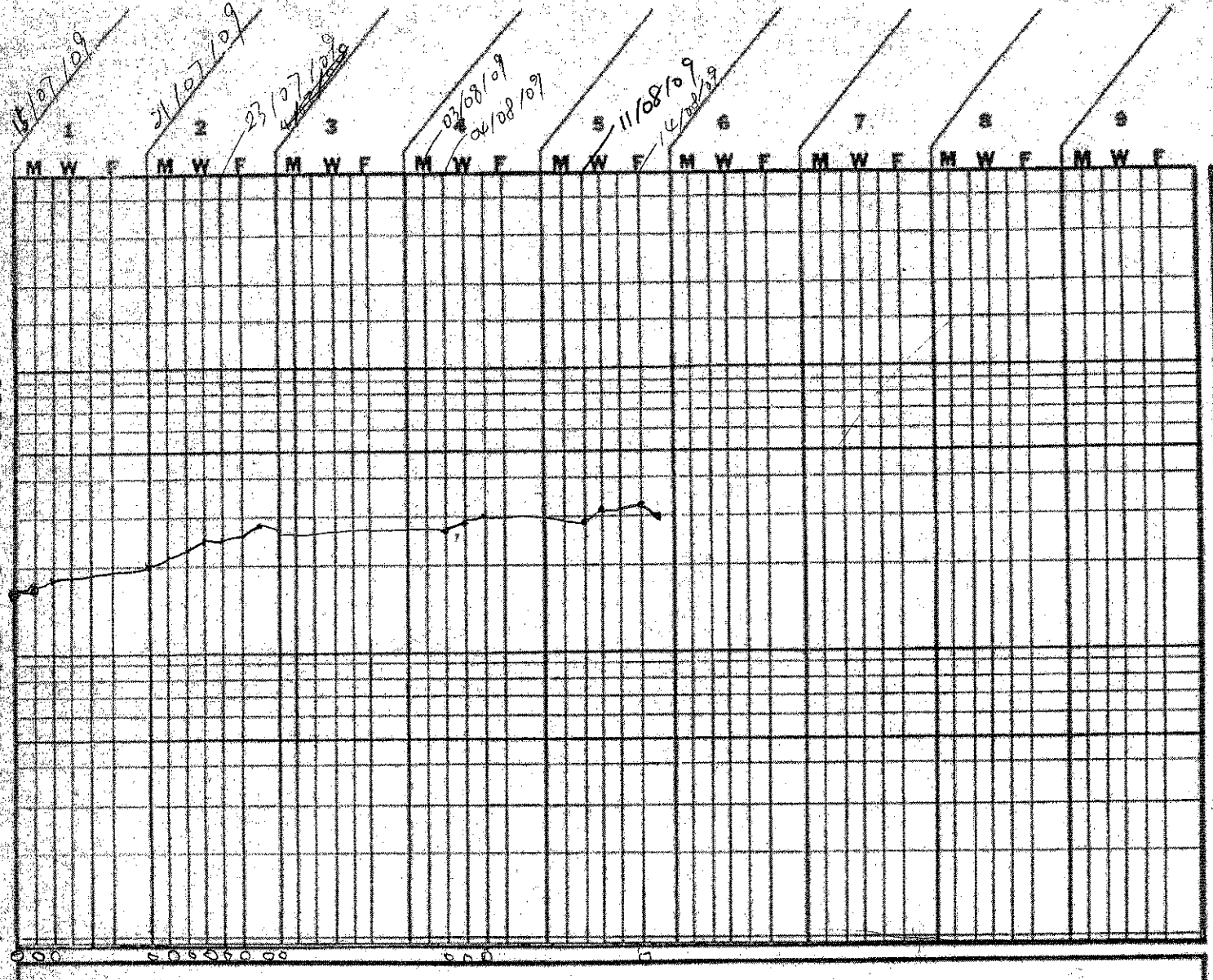


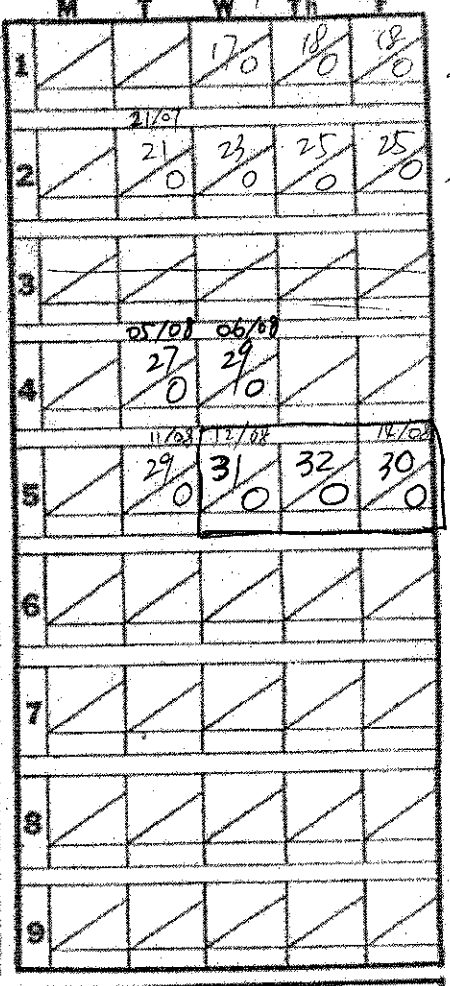
NAME _____ 3065 GRADE _____

BEHAVIOR Learning 20 Stats Terms GOAL 30/min

AC-4

BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13289
GAINESVILLE, FLORIDA 32604

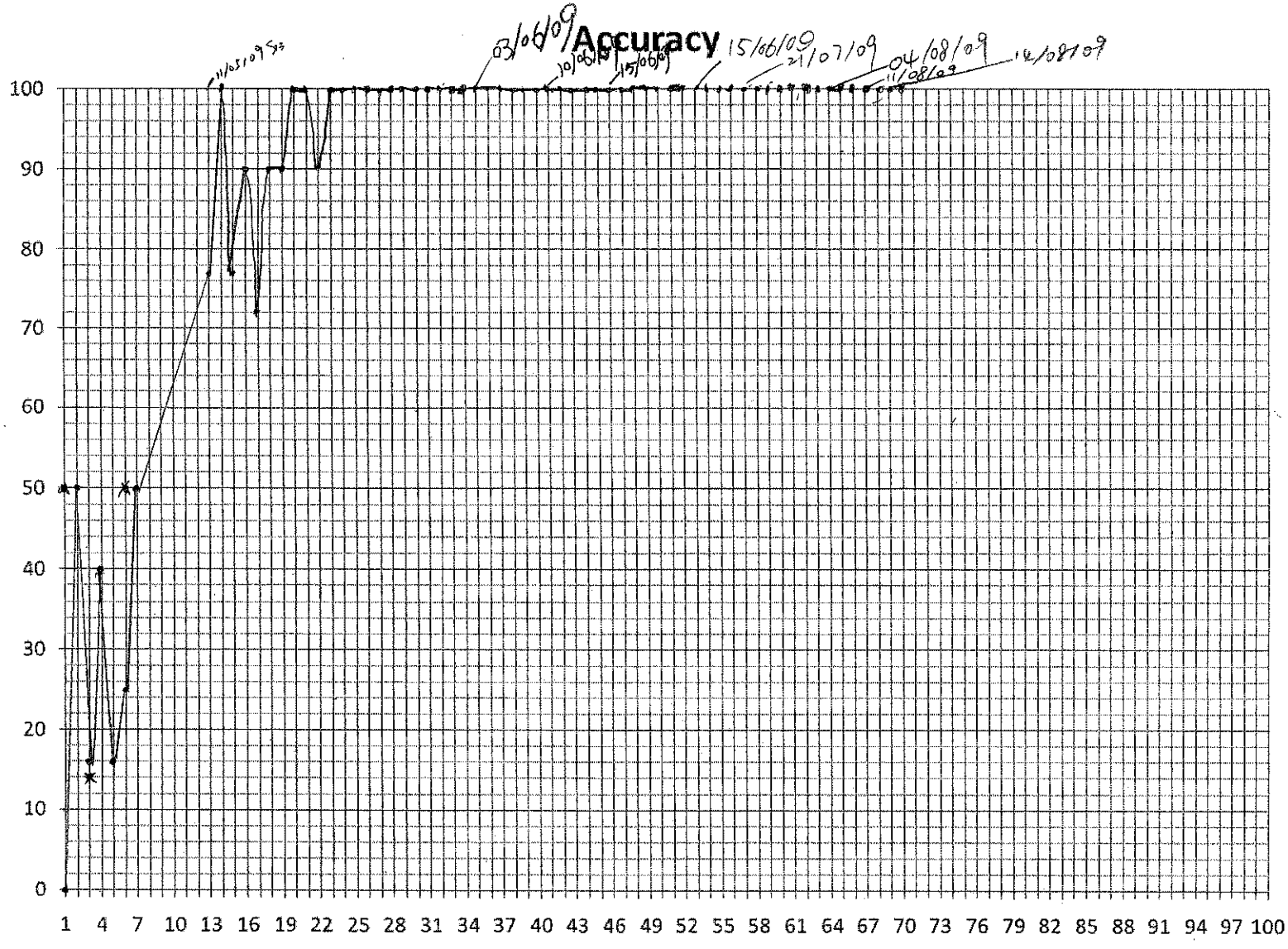




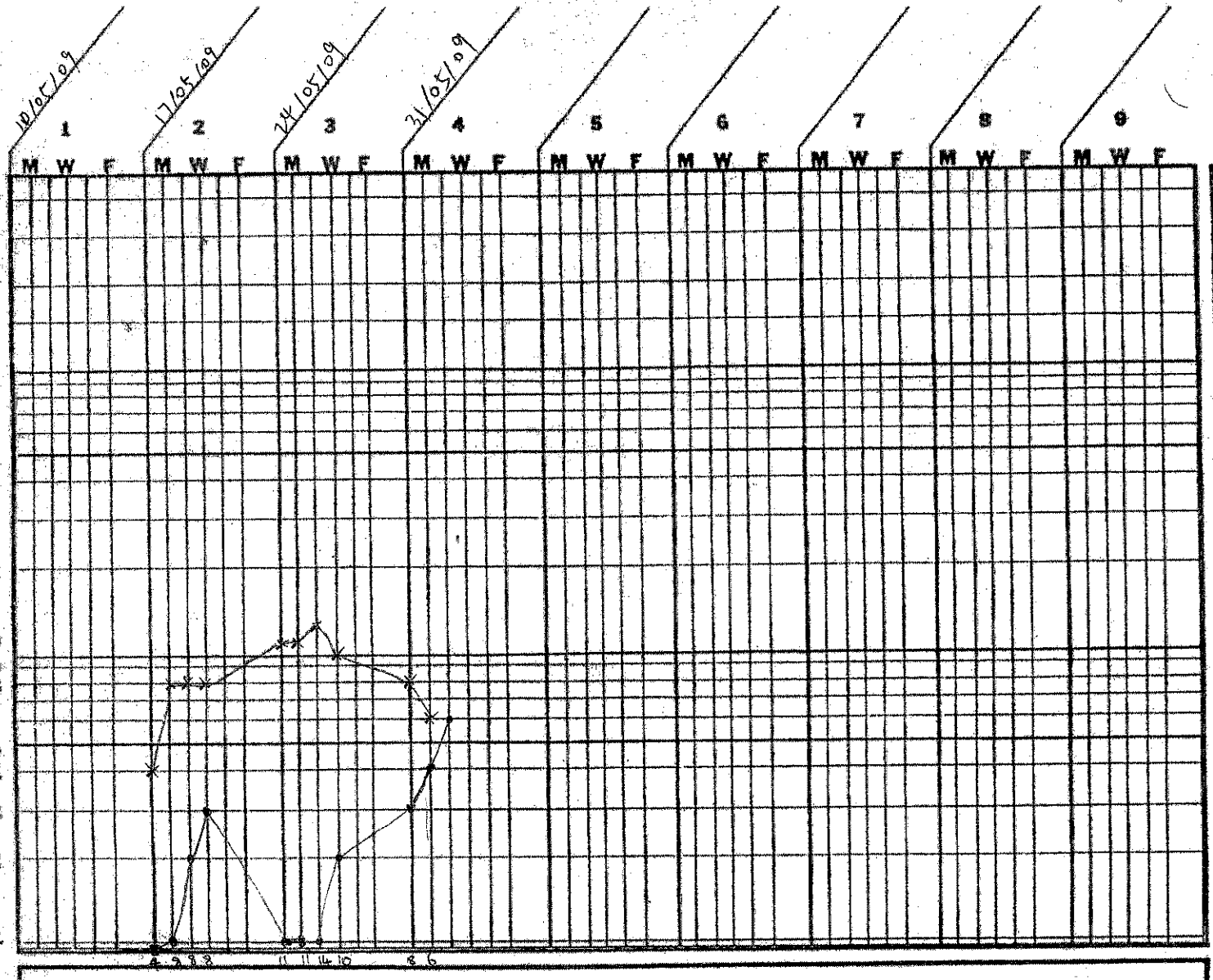
NAME (3065 GRADE

BEHAVIOR Learning ~~30~~ 20 Stats Definition GOAL 30/min

Participant 2065 date started date ended



AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13269
GAINESVILLE, FLORIDA 32604

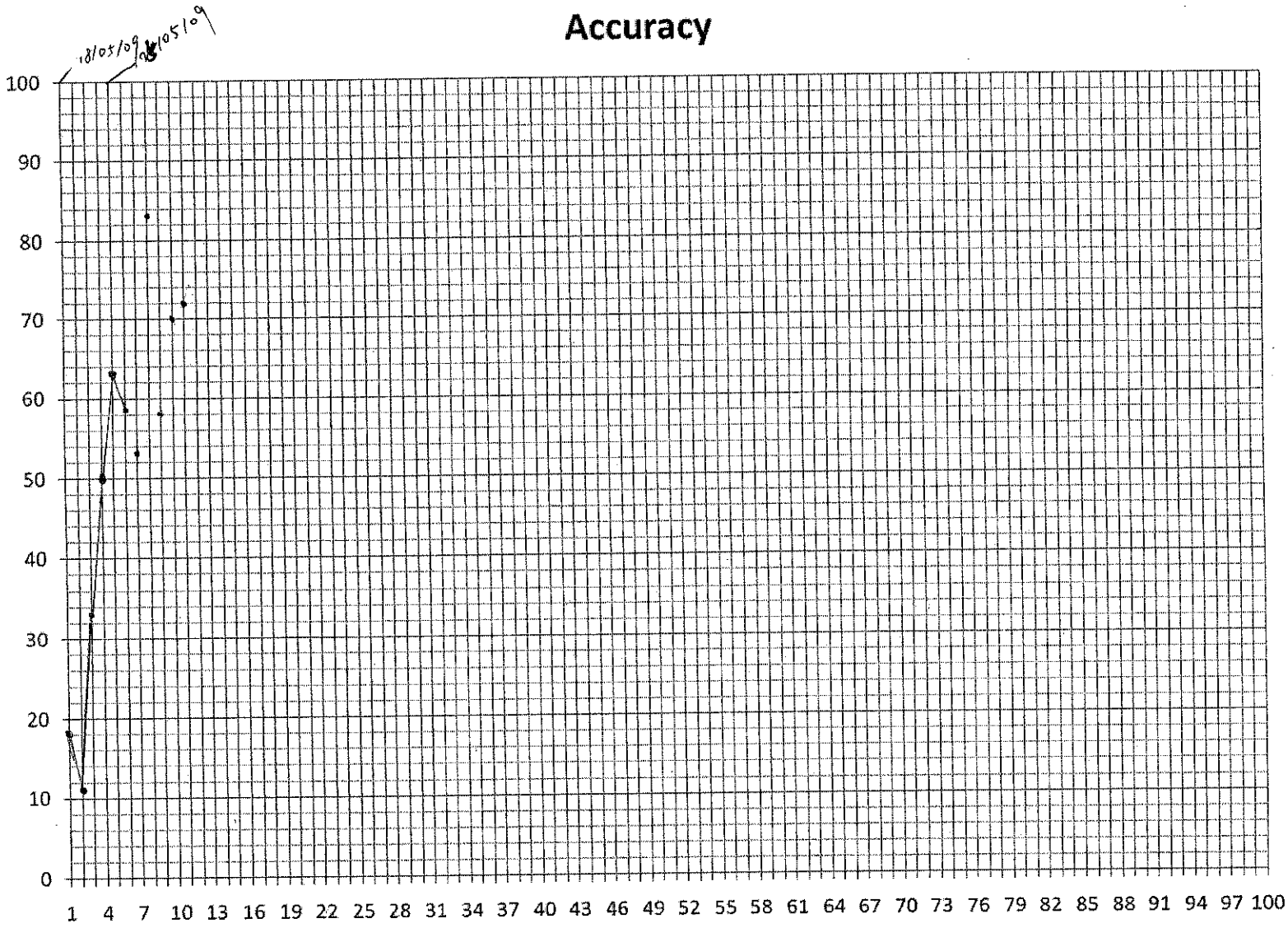


	M	T	W	Th	F
	18/5	19	20	21	
1	0/14	1/8	2/8	3/8	
	25	26	27	28	
2	1/11	1/11	1/14	2/10	
	02	03	03	06	05
3			3/8	4/6	6/4
4					
5					
6					
7					
8					
9					

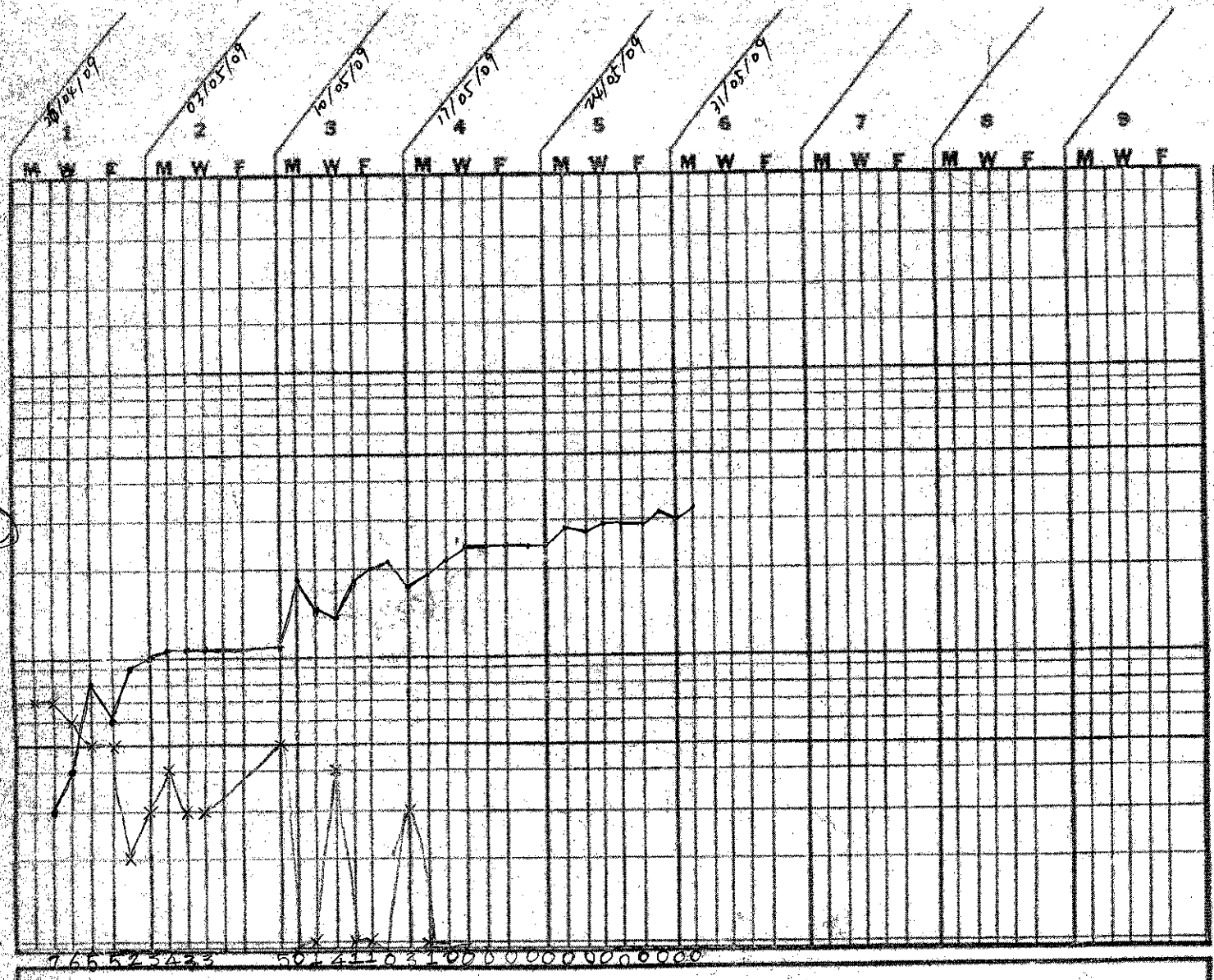
NAME 3075 GRADE

BEHAVIOR Learning to Stats Definitions GOAL 30%/min

Participant 3075 date started 12/05/09 date ended



AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13789
GAINESVILLE, FLORIDA 32604

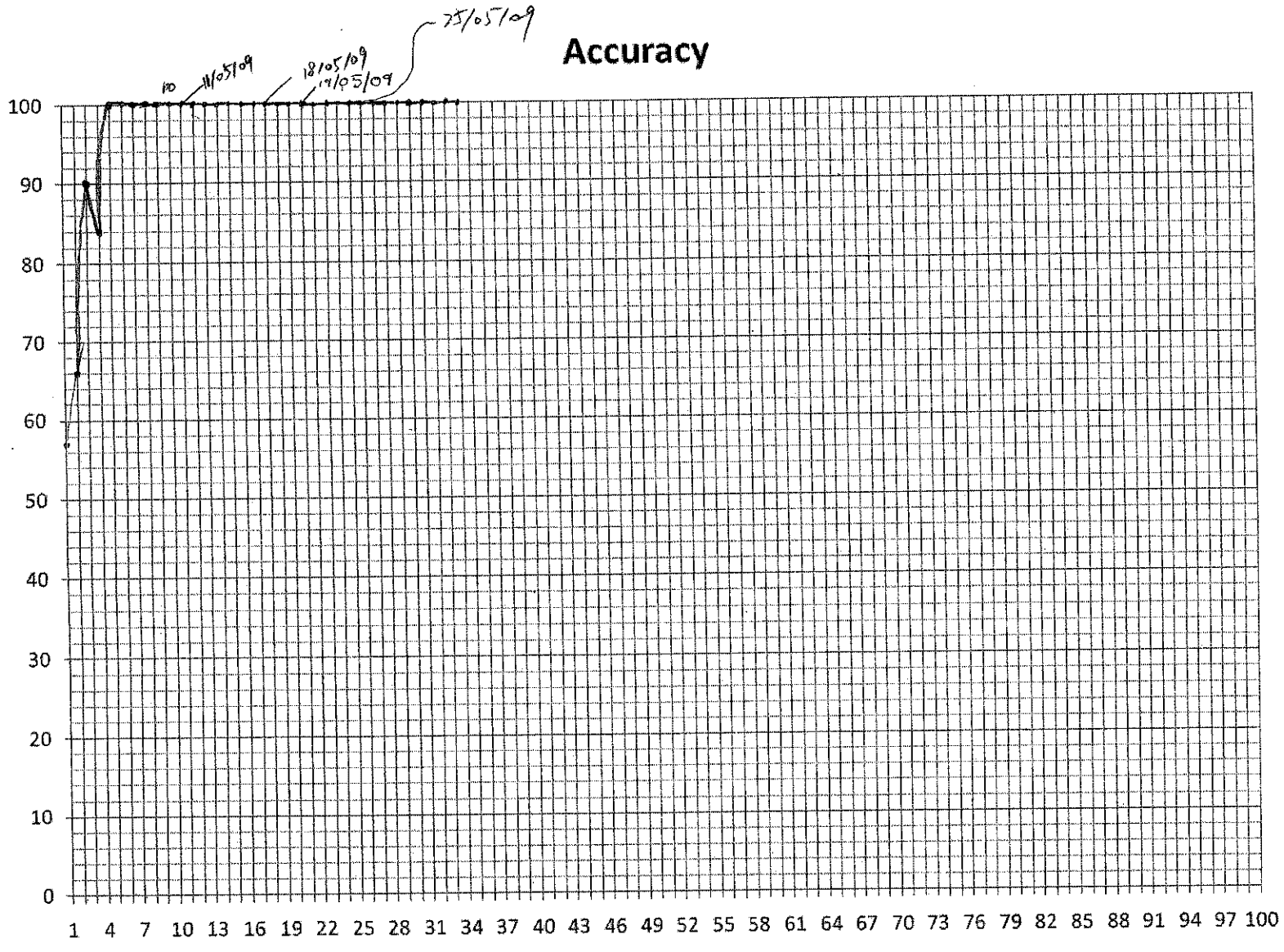


	M 27/04	T 28/04	W 29/04	Th	F 01/05	02/05
1	32/7	5/7	8/5	06/7	7/6	9/2
	05/05	06	07	08		
2		10/3	11/4	11/3	11/3	
	11/05	12/05	13/05	14	15	16 17
3	11/5	18/0	15/1	14/14	17/1	20/1 21/0
	18/05	19		21/05	22	23/05 24
4	17/3	19/1	22/0	24/0	25/0	25/0 2
	26/05		27		28/05	31
5	25/0	28/0	29/0	29/0	31/0	30/0 3
6						
7						
8						
9						

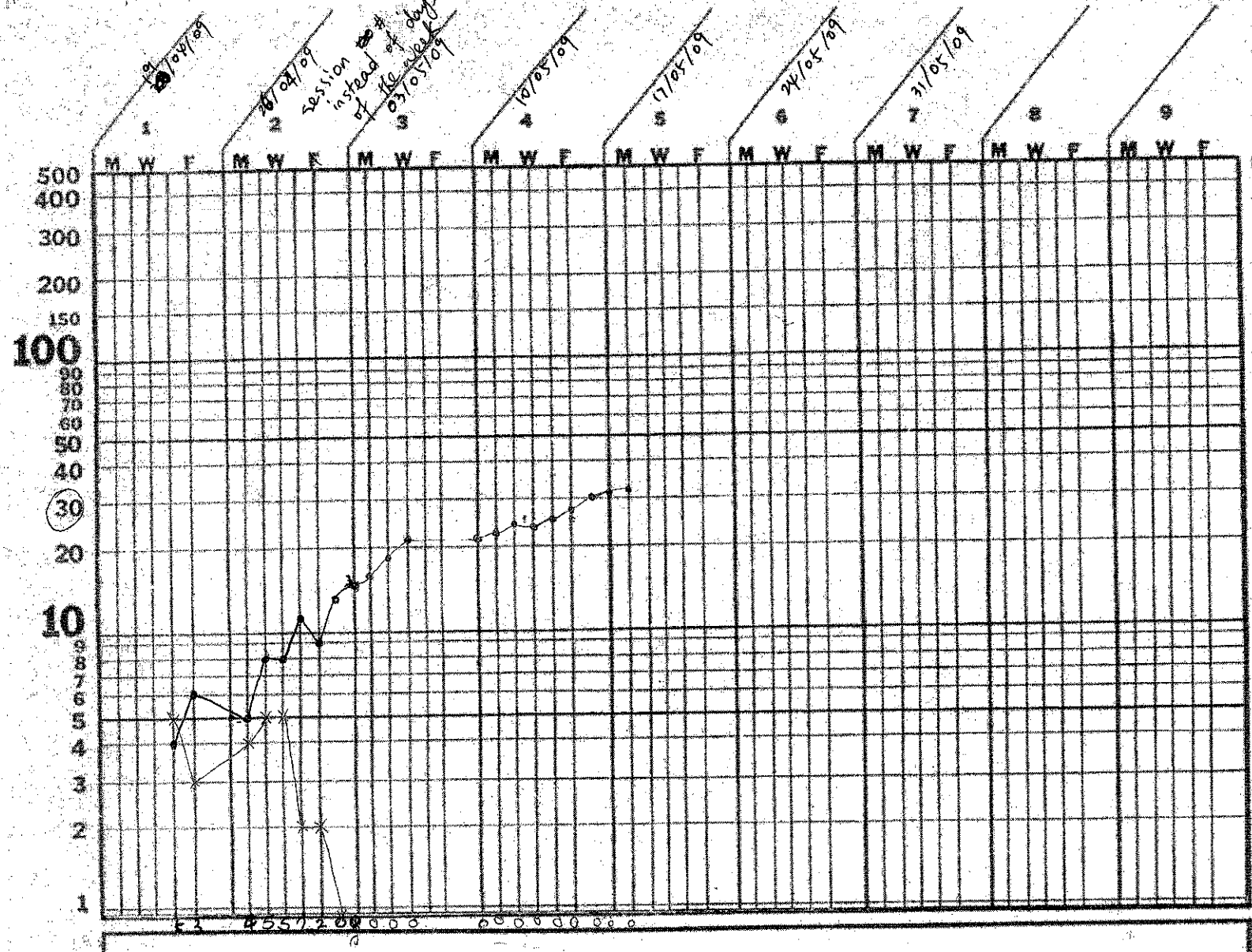
NAME _____ e 3074 GRADE _____

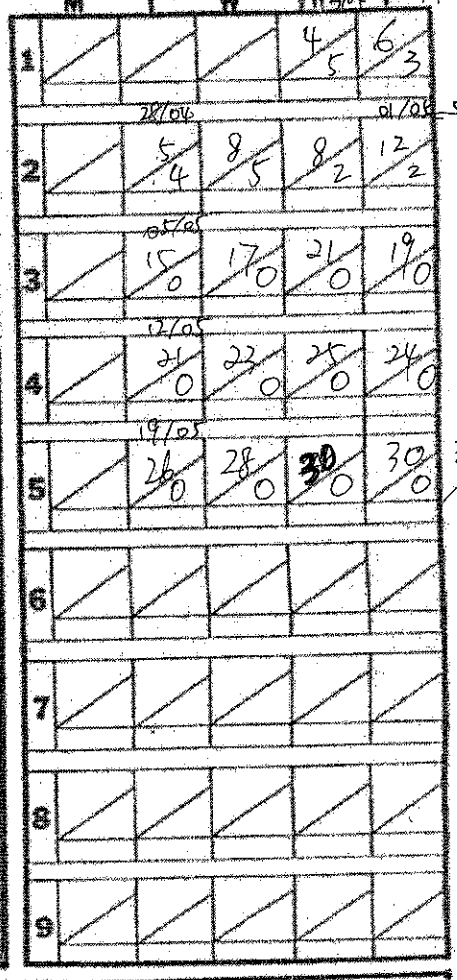
BEHAVIOR Learning 20 Stats Definitions GOAL 30/min

Participant 3074 date started date ended



AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13289
GAINESVILLE, FLORIDA 32604

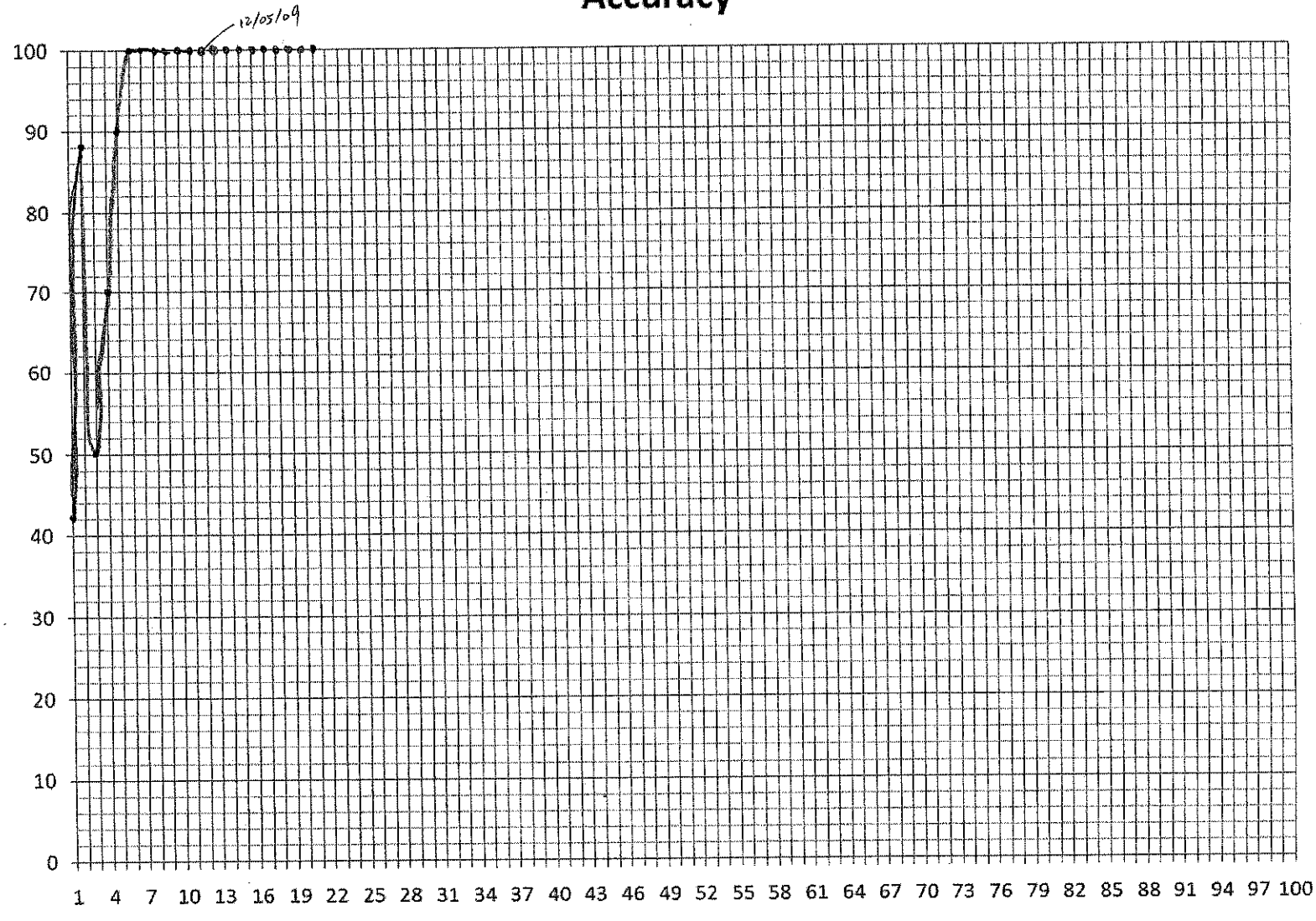




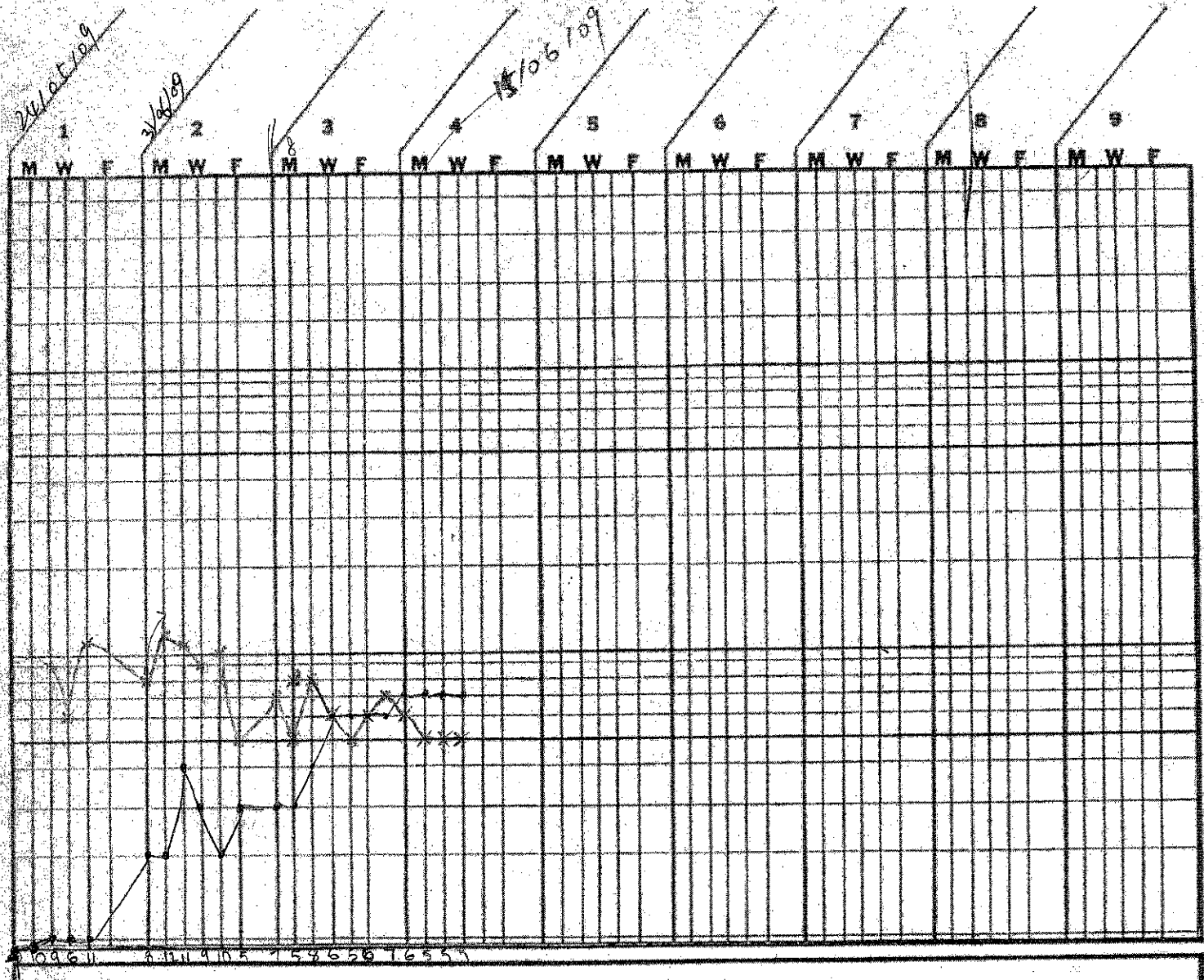
NAME 3071 GRADE

BEHAVIOR Learning Stats Terms (20) GOAL 30/min

Participant 3071 ate started date ended



AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13269
GAINESVILLE, FLORIDA 32604



	M	T	W	Th	F
1	0/7	0/10	1/9	1/6	1/11
2	2/8	4/11	3/9	2/10	3/5
3	3/7	4/8	6/6	6/7	7/6
4	7/5	6/5	7/5	7/5	
5					
6					
7					
8					
9					

NAME 3069

GRADE

BEHAVIOR Learning 20 STATS Definitions

GOAL 30/min

Participant 3069 ate started 24/05/09 date ended