ASSESSING PSYCHOLOGY MAJORS' KNOWLEDGE OF PSYCHOLOGY

USING THE AUBURN PSYCHOLOGY TERM TEST (APTT)

Except where reference is made to the work of others, the work described in this thesis is my own or was done in collaboration with my advisory committee. This thesis does not include proprietary or classified information.

_____

Julie A. Brandt

Certificate of Approval:

_____	_____
Jeffrey S. Katz Associate Professor Psychology	Lewis Barker, chair Professor Psychology
_____	_____
Alejandro Lazarte Assistant Professor Psychology	Joe F. Pittman Interim Dean Graduate School

ASSESSING PSYCHOLOGY MAJORS' KNOWLEDGE OF PSYCHOLOGY USING THE AUBURN PSYCHOLOGY TERM TEST (APTT)

Julie A. Brandt

A Thesis

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Master of Science

Auburn, Alabama
December 15, 2006

ASSESSING PSYCHOLOGY MAJORS' KNOWLEDGE OF PSYCHOLOGY

USING THE AUBURN PSYCHOLOGY TERM TEST (APTT)

Julie A. Brandt

_____
Signature of Author

_____
Date of Graduation

THESIS ABSTRACT

ASSESSING PSYCHOLOGY MAJORS' KNOWLEDGE OF PSYCHOLOGY

USING THE AUBURN PSYCHOLOGY TERM TEST (APTT)

Julie A. Brandt

Master of Science, December 15, 2006
(B.A. Elizabethtown College, 2002)

62 Typed Pages

Directed by Lewis Barker

Student assessment is a pressing issue for educators. This research addresses the proposition that students who perform better in psychology classes do so because they learn psychological terms and concepts with greater precision. Forty-two participants, graduating seniors from 2 different psychology programs, completed the Auburn Psychology Term Test (APTT), a yes-no recognition test of psychology key terms. The Approaches to Studying Inventory (ASI), ACT, GPA, and written EVAψ (Ecological

Validity Assessment of Psychological knowledge) were also collected to address the issue of construct validity. The ASI provides a measure of study habits via 2 variables: Reproduction and Meaning. A student with a Reproduction orientation to studying prefers clear-cut standards about the material they are required to learn, so that they may perform successfully on tests and achieve the grade or the degree they seek. Although not mutually exclusive, the Meaning orientation describes a student who approaches the educational process in order to use new knowledge in acquiring a deeper and more integrated understanding of his or her world. No differences between the 2 sample populations were found. Other important findings include APTT's lack of correlation with psychology GPA, which reflects its usefulness as a tool that assesses something unique about the knowledge graduating seniors have obtained. APTT scores predicted greater EVAψ scores, indicating greater quantity and quality of written application of psychology terms to psychological phenomenon in a cartoon. This provides more support for the assertion that the APTT is a highly useful assessment tool – performance on it correlates with performance on tasks that address all 3 of American Psychological Association's guidelines for psychology major assessment. Additionally, the ability to correctly reject pseudo-terms is key to better performance on the APTT, and this ability correlated with higher ACT and lower Reproduction orientations to studying.

Style Manual Used:

Publication Manual of the American Psychological Association

Computer Software Used:

Microsoft Word
SPSS
Minitab
Excel

TABLE OF CONTENTS

INTRODUCTION

The words comprising language grant their users the capacity for expansion and flexibility in information processing. Learning the appropriate word labels for objects, people, and concepts is likely the foundation upon which effortless thinking, writing, and speaking about a domain of information is built. Education within a given field should consist, at least partly, of accurately transmitting the word labels used by the shared community of language users. Regardless of which memory model currently dominates, there does exist a consensus about limits to information processing and to the overall function of human memory. By constructing for students as many "shorthand notations" of complex concepts as possible (a process called "chunking" by memory theorists), the educator provides a situation whereby thought processes are more likely utilized to form new, transformed or combined ideas and formulations (see Lindsay & Norman, 1977, p. 483). Word labels (or key terms, in the current research), can be thought of as chunks.

Education in the domain of psychology is unique in that it overlaps with the kind of learning and thinking that people have engaged in almost since birth. The development of reasoning about goals, perceptions, beliefs, and emotions begins between the ages of one and two, and is considered sophisticated enough to incorporate complexities of the human experience like pride and greed by age 14. The theory of mind, as it is often referred to, is a human capacity that appears and is honed throughout the early years of life. But the scholarly study of psychology requires that its students learn terms and

concepts which may sometimes augment, but sometimes compete with, the "the basic architecture of the system for understanding other minds" (see Saxe, Carey, & Kanwisher, 2004, for a review). Perhaps the following quote helps capture the idea: "Unlike behaviorists, normal adults attribute to one another (and to themselves) unobservable internal mental states, such as goals, thoughts, and feelings, and use these to explain and predict behavior" (Saxe, Carey, & Kanwisher, 2004, p. 88). In some ways, all human beings practice an informal "science of psychology" every day of their lives.

However, there is an established domain of scientific psychology, and many students attempt to familiarize themselves with it, by majoring in psychology and completing the requirements for a degree. This research provides an examination of the knowledge that graduating psychology majors have gained, and the individual difference variables that impact their ability to recognize terms that they have been exposed to as part of their coursework.

APA Taskforce

The American Psychological Association recently addressed the need for assessment of psychology majors' achievement (Halonen, et al., 2002). They assert that the "knowledge base" of psychology is too complex for educators to rely completely upon traditional assessment methods (i.e. grade point average, in-class examinations). According to the APA's task force, locally-developed achievement tests for graduates should assess whether psychology majors can: 1) demonstrate familiarity with the major concepts, theoretical perspectives, empirical findings, and historical trends in psychology, 2) demonstrate knowledge and understanding representing appropriate breadth and depth in selected content areas of psychology, and 3) use the concepts, language, and major

theories of the discipline to account for psychological phenomena (Halonen, et al., 2002). The present research represents an attempt to assess psychology majors according to these guidelines. The main testing method, a yes-no recognition test for key terms in psychology, is hypothesized to provide evidence for (1) and (2). The third guideline for student assessment, use of language to account for psychological phenomenon, was examined via a written response to popular media (a cartoon). This variable will henceforth be referred to as the written EVAψ (Ecological Validity Assessment of Psychological knowledge).

Print Exposure and Other Yes-No Measurements

Yes-no recognition tests are commonly used in research addressing issues of knowledge and literacy. Specifically, Stanovich and colleagues have rigorously tested the proposition that recognition of sources of information (print exposure) correlates with tests of declarative knowledge (West & Stanovich, 1991; Echols, Stanovich, West, & Zehr, 1996; Siddiqui, West, & Stanovich, 1998). The yes-no methodology provides a technique with which to circumvent the social demand confound of exaggerating or guessing at the number of recognized sources of information, by embedding a list of real author and magazine names within a larger list containing bogus names. The print exposure test emerged as a significant and enduring predictor of performance on tests of declarative knowledge – vocabulary, reading comprehension, general information, and spelling (Echols, Stanovich, West, & Zehr, 1996; Siddiqui, West, & Stanovich, 1998).

Particularly relevant to the present research is the fact that West and Stanovich (1991) also measured psychology content knowledge. Their test was used as part of the assessment curriculum for graduating seniors, which mainly consisted of 48 multiple-

choice items. This test of psychology domain knowledge emphasized experimental design, developmental psychology, statistics, and history and systems. The researchers also expanded the print exposure methodology to include six measures of media exposure – Newspaper, Magazine, Author, Television Show, Television Character, and Popular Film – but performance on this portion of the assessment contributed nothing to the regression model explaining variability in psychology-domain-specific knowledge. Only SAT scores and GPA predicted with statistical significance individual differences in performance on the psychology multiple-choice items. Essentially, reading and/or immersion in the non-specific domain of popular culture had no impact on the domain-specific knowledge of psychology majors (as measured by a multiple-choice test).

The second language learning literature has guided much of the thinking about this research. Although distinctions can be made between the acquisition of a domain of academic vocabulary and of a foreign language (Nation, 2001, p. 187), the similarities seem even more relevant. Huibregtse, Admiraal, and Meara (2002) refer to second language word knowledge as "not all or nothing, but rather gradual," and claim that much "sophisticated guessing" can be assumed to occur (p. 230). For this reason, they assert that the yes-no test is appropriate for measuring the size of passive vocabulary of second language learners. Additionally, Mochida and Harrington (2006) report that the yes-no test scores are strong predictors of test performance when words are ranked according to frequency of occurrence.

Signal Detection and Familiarity

The specific yes-no recognition test utilized for the present research, the Auburn Psychology Term Test (APTT), prompts psychology students to designate whether they

know whether a term is real or fabricated, based on a judgment of whether they were familiar with it as a term they had previously studied. Signal detection theory has been applied to memory research using much shorter lists of to-be-remembered words and much less diffuse study events (i.e. DRM false memory paradigm; Roediger, H.L. & McDermott, K.B., 1995), but the general theoretical framework of overlapping distributions of familiar vs. novel stimuli applies quite well to the APTT. For instance, despite the differences between testing recognition of items from a single encoding event and from 4 years of education, the following statement adequately describes the context that most likely gives rise to APTT items that are correctly accepted as psychology terms (hits) – "A moderate amount of recollection coupled with a reasonably high level of familiarity would point strongly to the prior occurrence of the item on the list (more strongly than would either process considered alone)" (Wixted & Stretch, 2004, p. 633).

Recollection, in this quote, refers to the presence of episodic memory for the specific encoding event(s). For psychology majors learning key terms, recollection implies that the student remembers encountering the term in a textbook, in a lecture, in their notes, or perhaps in a journal article. Familiarity, on the other hand, is influenced by how frequently a term has been seen (MacMillan & Creelman, 2005, p. 122). If the term were linked many times (with multiple applied examples, or to many related terms or theorists) to a pre-existing conceptual framework, the student would be assumed to have achieved a maximum level of familiarity. Combining recollection of specific study events with an overall level of familiarity due to repeated exposure across 4 years of coursework is presumed to provide a student with enough recollective data to designate that "yes, that is a term I have studied" (classified as a hit, in signal detection theory), or,

perhaps more importantly, "no, that is not the same as any terms that I have studied" (classified as a correct rejection). Incorrect responses to the real terms are designated misses, and incorrect responses to the fabricated terms are false alarms (Tanner & Swets, 1954; Coombs, Dawes, & Tversky, 1970; White, & Wixted, 1999).

Levels of Word Knowledge

A central question posed by this research – how do students come to reject the fabricated terms? – can be addressed using Baumann and Kameenui's (1991) levels of declarative vocabulary knowledge. The shallowest level is association, when the student can link the term with a single definition or context. The intermediary level is comprehension, which is apparent when the student can produce an antonym, or demonstrate understanding of the word in multiple contexts, or classify the word into a category. The deepest level of processing is reflected by generative knowledge, whereby the student can produce novel responses to the word (original sentences), or can restate the definition in their own words. Note that there is a high degree of similarity between these 3 levels and the aforementioned APA recommendations for psychology major assessment.

Presumably, the deeper levels of vocabulary knowledge are accompanied by a) multiple links to other real terms in a semantic network, b) numerous encounters with the correct spelling of the term, and c) several contexts in which the student heard the term pronounced by professors, other students, or herself. After 4 years of coursework, many students can be expected to have progressed to the second and third levels of vocabulary knowledge. Therefore, attempts to process the semantically similar fabricated terms would be accompanied by enough conflicting information to reject them. These are most

likely the conditions under which students who can reliably reject foils in favor of real terms are produced. These same conditions are hypothesized to produce students who are capable of using (writing) psychology's key terms to account for psychological phenomenon.

Approaches to Studying Inventory

McCabe, Presmanes, Robertson, and Smith (2004) found that, for 10-word lists of studied items, instructing participants to engage in item-specific processing rather than relational processing resulted in fewer false alarms for critical lures. They cued item-specific processing with the phrase "think of a unique characteristic" of each studied word, and cued relational processing by asking participants to focus on the similarities among the studied words, and to try to relate the words to one another (p. 1076). This prompts the question, in reference to the present research – do particular study habits encourage better remembering (less false alarms)?

Several methods for assessing student study habits were considered, but due to the social desirability confound for self-report of study time, as well as the variability in students' efficient use of study time, an 18-item questionnaire meant to characterize preferred objectives for studying, rather than quantify study time, was used (for the original 64-item Approaches to Studying Inventory (ASI), see Entwistle & Ramsden, 1983). Newstead (1992) administered the short form to 188 participants studying for an honors psychology degree. In discussing the short form's usefulness, he concluded that the 18-item ASI was moderately reliable across the 3-year span of the study (reliability scores on the three subscales range from $r = .44$ to $r = .61$), and that it demonstrated a reasonably robust factor structure (p. 318-20).

The ASI includes items that load on 2 factors: Meaning and Reproduction. The type of statement that meaning-oriented students answer favorably toward is "I spend a good deal of my spare time in finding out more about interesting topics which have been discussed in class." Reproduction-oriented students agree more markedly with statements such as "I tend to read very little beyond what's required for completing assignments" (Newstead, 1992, p. 301). As might be expected, the scores on the Meaning subscales of the ASI significantly correlated with the students' performance on two different years of exams, but scores on the Reproduction subscale showed no correlation with academic performance (Newstead, 1992, p. 320).

Students with a predominantly meaning-oriented approach to studying seek to integrate the material into their pre-existing domains of knowledge (a relational approach rather than an item-specific approach). It might be argued that students who are motivated to acquire a deep understanding of psychology, through study time devoted to extracting meaning, may be more likely to have concrete, easily discriminable representations of key terms. As well, students with a meaning orientation may have the confidence to definitively reject the fabricated terms, due to their recollection of studying psychology texts extensively. However, it might also be argued that students with a reproduction orientation to studying, who prefer clear-cut standards about the material they are required to learn (a more item-specific approach), benefit from the protection against false memory that McCabe and colleagues identified (2004).

Research Objectives

The APTT was developed in order to assess learning of the domain knowledge of psychology. A list of 100 terms comprising the APTT was generated by drawing 50

bolded or italicized terms of an introductory textbook (roughly the same number from each chapter). Fifty other "psychological sounding" terms comprise the pseudo-terms, or foils. The resulting 100 terms were randomly arranged, and students were instructed to discriminate between the two varieties of stimuli.

Hypothesis 1: Better performance on the APTT correlates with ability to use psychology's key terms to account for psychological phenomenon, in written responses to a short cartoon clip (EVA ψ).

Students' earlier forms of unsophisticated psychological reasoning and vocabulary (i.e. Saxe, Carey, & Kanwisher, 2004) may compete with the application of their domain-specific knowledge to psychological phenomenon. As well, it has been observed that written language is a "special network" within the already existing basic cognitive architecture, and that the semi-independence, or modularity, of second language written vocabulary is even greater than that of second language reading vocabulary (Donald, 2001, p. 304). However, despite the lack of perfect correspondence, Meara and Fitzpatrick (2000) report a high correlation between receptive vocabulary and productive vocabulary, particularly for generation of infrequent words.

In order to address the APA taskforce's third guideline under highly challenging conditions, it was deemed appropriate to provide students with a stimulus they were exposed to both prior to and outside of the realm of psychological learning. Students were asked to account for psychological phenomenon in a short clip from *The Simpsons*, which has been on television since 1989 (or since the graduating seniors were approximately 7 years old).

Hypothesis 2: APTT performance of Auburn psychology graduates does not differ from that of graduates at another institution.

Alternate forms of the APTT have produced a convincing degree of alternate form reliability (Smith, 2006, p. 32), providing support for the claim that the indication of learning provided by the APTT is due to the overall theoretical framework of increased word discriminability, and not due to the specific items on any particular version. Administering the APTT to graduating seniors at a small college in another region of the country provides further opportunity to address this claim. Colleges and universities can be expected to differ in any number of ways – the student body they attract, the educational theories they abide by, the dominant political orientation of the surrounding area, to name a few. However, the core of key terms and concepts in a field are maintained by communities of interacting professionals. Because there is a relatively constant domain of knowledge (see Boneau, 1990, for a list of key terms that were democratically ranked by psychology professors and textbook authors), it is assumed that student knowledge will not differ across institutions.

Hypothesis 3: Better performance on the APTT correlates with higher general cognitive ability, as approximated by ACT scores.

Hypothesis 4: Better performance on the APTT correlates with greater success in completing coursework, as indicated by grade point average (GPA).

Hypothesis 5: APTT performance is positively related to performance on a yes-no recognition test for sources of psychological information (journals), and to scores on a measure (derived from questions on the ASI) of psychology-related reading.

Stanovich and his colleagues reported that both SAT and GPA predicted multiple choice psychology test performance among graduating seniors (West & Stanovich, 1991). These findings are the basis for Hypotheses 3 and 4. They also found that measures of print exposure (newspaper, magazine, and popular author), which they claim indicate amount of out-of-school reading, did not relate to psychology test performance (1991). This finding provides the rationale for Hypothesis 5.

Hypothesis 6: APTT performance is related positively to the meaning orientation and negatively to the reproduction orientation to studying, as measured by the ASI.

Although a more item-specific, reproduction-oriented student may enjoy the short-term benefits of learning small lists of words and accurately recognizing them on tests following short delays, it is hypothesized that this approach actually hinders the student's long-term learning. Furthermore, the relational, meaning-oriented approach to studying is hypothesized to promote the formation and persistence of greater discriminability of key terms.

Hypothesis 7: Graduating psychology majors' performance on the APTT is a strong indicator of domain-specific learning that persists beyond the time(s) that it was measured by in-class, course-specific assessment.

Pilot work has shown that students tested in the first few days of an introductory psychology class perform no better than chance. However, students tested again in the final days of the class are much better able to discriminate between true psychology terms and fabricated terms (reported in Smith, 2006). Indeed, the better performance on the post-tests might reflect a pedagogical achievement: psychology courses provide students with domain-specific information. Furthermore, psychology student knowledge may

extend beyond the "common sense" pop psychology that so regularly enters the wider lexicon of our culture. Since pre- and post-testing students who take psychology courses has established that performance on the APTT is a strong indicator of learning that is bounded by an interval of a few months, the present research examines APTT's broader usefulness in student assessment.

METHOD

Participants

As part of the psychology department's assessment of graduating seniors, students were recruited to complete the APTT and an additional questionnaire. They were assured that the data gathered for the purposes of the study was not a requirement for graduation, nor would it affect their standing with the department. A total of 31 students (out of 111 graduating seniors) participated in the research. Additional participants were recruited from Elizabethtown College. Of the 14 psychology majors scheduled to graduate in Spring 2006, 11 completed the APTT and the ASI.

Materials

APTT: The list of 50 real psychology terms, randomly arranged among 50 foils, appears in 3 columns on a single page (see Appendix A for the version used, and Appendix B for a version containing the true terms in bold). These instructions appear above the list:

Below, 100 terms are listed. Some of them are key Psychological terms that you encountered in lectures and reading the textbook. Others will be unfamiliar to you, because they are bogus, fabricated terms that sound like psychological terms, but are not "real" psychology terms. Please look at each item, then bubble in "A" if you recognize it as a real term, and "B" if you think the term is bogus.

The following instructions also appeared on the informed consent document, as a preliminary explanation of the APTT:

Your task is to identify which of the terms are real and which are fabricated. *For example, terms such as* ***memory*** *and* ***Ivan Pavlov*** *are both associated with psychology, so you would mark "A" on the scantron. Likewise,* ***intestinal myopia*** *and* ***terminal distress*** *are not part of psychology, so for those items you would mark "B".*

ASI: The 18-item Approaches to Studying Inventory (ASI; Newstead, 1992; see Appendix C), included the following instructions: *Please answer the following questions, mostly about your study habits over the past several years, and your academic plans and interests, honestly and objectively. Circle a number to indicate the extent to which you agree or disagree with these statements.* The numerals 1, 2, 3, 4, and 5 are typed to the right of each statement, with the words "Disagree," "Neutral," and "Agree" appearing above the 1, 3, and 5, respectively. Therefore, a response of 5 indicated maximum agreement.

The single, one-sided sheet of paper, the Scantron scoring sheet, and the additional questionnaire (see Appendix C) were identical for both samples of graduating seniors. Additional information was gathered from Auburn students, including a test of familiarity with psychology journal titles (see Table 1 for true and foil journal titles used) and a writing sample.

EVA ψ: The students were given space on a single sheet of paper, below the following prompt, on which to record responses: ***"Brainstorm" from what you have learned from your course work:*** *Watch the following* Simpsons *clip (it's only about 10 minutes). Write down any psychological concepts you can find (even loosely) illustrated.*

You might include how certain concepts or theories are supported by things that occur. You could also write about how the cartoon presents misconceptions of violations of psychological concepts. There are no "wrong" answers, so feel free to be creative"

The clip was a short vignette from the fifth "Treehouse of Horror," episode 109. It was based on a Ray Bradbury story called "The Sound of Thunder," from which the term "butterfly effect" (referring to the chaos theory concept of sensitivity to initial conditions) was originally taken.

Procedure

Students were mailed a letter by the Auburn University department chair, and contacted by emails from the researcher to set up specific times to participate in the department's assessment research. At Elizabethtown College, students were contacted by emails from one of the professors, and invited to participate in research being conducted by a previous graduate of the psychology program. Experimental sessions were sometimes conducted in groups (maximum of 8 at a time), and sometimes individually, depending upon student availability. Upon arrival, participants read and signed the informed consent form, and then were presented with the APTT, a Scantron scoring sheet, and the ASI. After students completed the APTT, a "debriefing" key was shown to them (individually or projected on the screen in front of the group), on which only true terms were bold (Appendix B). Auburn students also completed the additional measures of career interests and journal recognition, and were shown the 10-minute cartoon. They responded to the prompt, helped themselves to snacks and beverages, were thanked for their time, and dismissed. Some of the students were recruited by offering monetary compensation. They were paid $10 cash before the session began.

The difference between standardized scores for the proportions of correctly identified terms (hits) and incorrectly identified bogus terms (false alarms), or d', was computed. Because this statistic provides a way to account for liberal or conservative response tendencies, it was calculated as the best possible overall indicator of the status of the students' psychological vocabulary. Journal recognition patterns were also translated into d' scores. The 10 real journals were chosen from a list of APA journals, and the foils were concocted to be stylistically similar to additional real journals (see Table 1). The instructions read: *Below is a list of Psychology Journal Titles – again, some are real and some are not. Put a check-mark beside any that you recognize as actual psychology journals.*

ACT and GPA: Estimates of both overall and psychology grade point average were requested. Response rate among students who completed the research was 100%.

Career interests: Students responded to the following prompt: "Please circle any of the following types of careers that you have considered pursuing: Academic/Teaching, Clinical, Experimental, Industrial/Organizational, Other". Two additional questions were included: "Have you ever sought any kind of counseling? (optional)," with possible choices "yes" or "no," and "Do you plan to go to graduate school in psychology?" (yes, maybe, or no were possible choices).

Study Habits: Scores for 2 types of study habit orientations – Meaning and Reproduction, as well as a score assessing to what degree students engage in deep reading of class material (the Reading score) were derived from responses to the ASI. The factor structure that was hypothesized to exist produced less-than-convincing support in both a sample of pre-medical students (Newstead, 1992, p. 320) and in our own

sample. For this reason, only the items that produced the strongest loading on the factor it was associated with were included.

A Reproduction orientation score representing a summation of the response values for the following 5 statements was computed:

> "I like to be told precisely what to do in essays or other set work," "I suppose I'm more interested in the qualifications I'll get than in the courses I'm taking," "Often I find I have to read things without having a chance to really understand them," "I tend to read very little beyond what's required for completing assignments," and "I find I have to concentrate on memorizing a good deal of what we have to learn."

A Meaning orientation score representing a summation of values for these 5 statements was computed:

> "I find academic topics so interesting, I should like to continue with them after I finish this course;" "I spend a good deal of my spare time in finding out more about interesting topics which have been discussed in class," "I often find myself questioning things I hear in lessons or read in books," "My main reason for being here is so that I can learn more about the subjects which really interest me," and "It's important to me to do really well in the courses here."

A Reading factor was developed by locating all questions that included a reference to reading. It includes two items that are part of the Reproduction score (so the statistical significance of their correlation cannot be interpreted as actually significant). The numeric responses to the following 2 questions –

> "Often I find I have to read things without having a chance to really understand

them" and "I tend to read very little beyond what's required for completing assignments,"

were subtracted from values for these questions –

"I usually set out to understand thoroughly the meaning of what I am asked to read," and "When I'm reading I try to memorise important facts which may come in useful later."

EVA ψ Scoring: The 4 criteria used to assign scores to the written responses include: 1) frequency count of key terms, *k* ; 2) frequency count of psychological concepts fully described, but lacking the appropriate label, *c* ; 3) frequency count of key terms accompanied by complete, accurate explanations, *b* 4) an additional count of "good points" that were extraneous to the domain knowledge of psychology, but still were relevant to the task of thoughtful analysis, *g*.

A total score was computed with the following equation: $T = k + c + g + 2(b)$. The rationale for weighting the variable representing key terms plus full explanations lies in the fact that this task was designed to measure more than simple familiarity with key terms. It provides students the opportunity to demonstrate both "knowledge and understanding representing appropriate breadth and depth" and to "use the concepts, language, and major theories of the discipline to account for psychological phenomena" (Halonen, et al., 2002).

The following examples clarify the criterion used: "Homer kept messing with the *dependent variables* when he went back in the future, by killing the fish it changed his future & his life…… It was almost like he was *hallucinating* in some of the scences which w/out drugs are symptoms of disorder" (emphasis added) would receive the

following score: T = 2 + 1 + 0(g) + 0[2(b)] = 3. The student showed evidence of conceptual knowledge from abnormal psychology (1*c), and used 2 key terms (2*k). Had the student used the correct term *independent variable* instead, it would be scored as a *b* rather than a *k*, and therefore would have an additional point added to the total score. The following is an example of a response given a score of 1* *g*: "The cartoon clip included the phenomenon of time travel and portrayed Homer traveling back to the time of the dinosaurs using a broken toaster. This is not necessarily a psychological concept (it's more of a scientific one)…"

RESULTS

Occasionally, a student omitted a response (totaled less than 0.1% of the cases). However, every Auburn student completed the APTT in its entirety. The 6 APTT items that Elizabethtown students did not respond to were numerically acknowledged in calculations of response probabilities. No attempt to replace any of the missing data points was made.

Table 2 shows the results of the APTT test for the Auburn and Elizabethtown students. The differences between the two sample populations were analyzed using a series of one-way ANOVA tests. No statistically significant differences were found for any of the pertinent variables (hits, correct rejections, total scores, d-prime, GPA in psychology, total GPA, intentions to attend graduate school, ACT, reading, reproduction, and meaning study orientation scores). The proportion of Auburn students planning for graduate school (13/31) was similar to that of Elizabethtown students (4/11), although more Auburn students responded "maybe" (13 versus 2), and proportionately more Elizabethtown students were certain about not planning for graduate school (5 in each sample). Again, these differences were not extreme enough to consider the samples as originating from different populations. See Table 2 for additional descriptive statistics.

Some differences between the sample populations are worth noting, however. Response rates were highly divergent (79% versus 28%), despite the fact that students were contacted by the social psychology professor at the small college, and by the chair

of the department at the public university. Auburn students were even offered money ($10) to complete the research.

The modal Elizabethtown graduate's knowledge of psychology terms differed somewhat from the Auburn graduate's. In addition to the universally accepted terms (addressed below), all 32 Auburn students correctly identified the term fixation. All 11 Elizabethtown students correctly identified action potential, introspection, long term memory, longitudinal study, libido, negative feedback, standard deviation, general intelligence (g), and dependent variable. Elizabethtown students had more trouble rejecting certain pseudo-terms, like "operant encoding" (58% vs. 82% correct at Auburn) and "captive nerve ending" (64% vs. 71% correct at Auburn) and "somatic transmission" (45% vs. 61% at Auburn). All 11 rejected "Festinger-Maslow effect", "hypostasis", "duozygotic twins", "schema taking score (STS)", "distance IQ", "instinctual deprivation", "law of effort", "echoic inhibition", "transperceptible difference (TD)", "superstitious relaxation", "cropping response", "hapless motivation", "involutional study", and "transferential psychosurgery."

The following analyses were conducted only with the Auburn sample. The d-prime for journals correlated with the d-prime on the APTT, $r(31) = .363$, $p = .045$. Journal recognition did not correlate with any other major variable, such as GPA or intentions to go to graduate school. Table 1 displays percentage correct responding for the journal recognition task.

EVAψ scores not only correlated with APTT performance, but also with GPA and intention to attend graduate school. Specifically, the frequency of key terms accompanied by full explanations was positively related to APTT total score ($r = .438$, p

= .016), APTT d-prime (r = .410, p = .024), and overall GPA (r = .415, p = .023). The frequency of terms included in the response, but unaccompanied by satisfactory explanation of their relevance, correlated with intentions to attend graduate school, r = .468, p = .009. The total score for the writing task was positively related to number of APTT hits (r = .366, p = .047) and overall GPA (r = .366, p = .047).

Combined Analyses:

Because the Auburn and Elizabethtown samples did not differ, the samples were collapsed and (unless otherwise noted) the following results and discussion reflect analysis of this larger data set (n = 42). It was hypothesized that the general cognitive ability and the study behavior of students would be the main modulating variables determining differences in performance on the APTT. The results and correlation values for the measures used to assess these variables are concisely presented in Table 3.

ACT: Although self-reported, this measure may be considered roughly accurate, as the sample average (M = 24.67, SD = 4.03) was no higher than that of Auburn University students (M = 24; $t(26)$ = .86, p = .40). When only SAT scores were available (for 14% of the participants), they were converted to the equivalent ACT (see http://www.californiacolleges.edu/admissions/extras/uc).

ACT was the strongest predictor of APTT scores. Closer analysis of the data revealed a differential pattern for APTT performance with true and bogus psychology terms. Somewhat surprisingly, the ability to identify real psychological terms was not related to ACT, in this sample (r = .158, p = .413). Correctly rejecting foils appears to be the most important determiner of d' scores (r = .544, $p < .001$), and it is strongly related to general cognitive ability, as approximated by ACT scores (r = .511, p = .005).

GPA: The highest correlation was found between psychology GPA and overall GPA (r = .791, p < .001). Psychology GPA was also strongly related to ACT (r = .477, p = .009) and the reading score (r = .349, p = .023). Surprisingly, only overall GPA predicted higher d-prime scores on the psychology term test (r = .338, p = .029). As well, overall GPA was negatively related to the reproduction study orientation (r = -0.334, p = .030).

Reading Score: The reading score correlated positively with greater intention to go to graduate school (r = .447, p = .003), and with both measures of GPA (see Table 4). However, it did not relate directly to performance on the APTT.

Meaning Orientation: Students with higher ACT scores tended to have more of a meaning orientation to studying, compared to their peers. However, this factor did not relate to performance on the APTT, or intentions to go to graduate school (r = .089, p = .634). The differences between this factor and the reading score lies in Meaning being related to ACT (reading was not), and in Reading being a predictor of graduate school intentions (meaning was not).

Reproduction Orientation: Students who reported study habits geared toward simply reproducing information for the purpose of evaluation were less likely to reject bogus terms (r = -0.385, p = .012). Their d-prime scores tended to be lower as well (r = -0.383, p = .012), indicating a general lack of discrimination between true and bogus terms.

Additional Analyses: Figure 1 shows true terms percent correct (hits) and foils percent incorrect (false alarms) for the entire sample. The abscissa represents terms arranged in ascending order of performance. Figure 2 is simply presented as a closer approximation of the signal detection model that this research is based upon. The distributions of true terms and pseudo-terms overlap slightly, and the "familiarity" with true psychology terms

is represented as stronger or greater than the familiarity that produces false alarms. Figure 2 might be conceptualized as a picture of the overall status of the modal psychology graduate's domain-specific knowledge structure.

Table 4 includes all 50 foils, rank ordered according to the percent of correct rejections. The inverse of these numbers was used to plot the line graph for the foil distributions in Figures 1 and 2, because the "opposite" of familiarity was used to reject the pseudo-terms. Foil familiarity can be considered familiarity with correct terms that are conceptually similar (i.e. "neuroresponder" is *not* familiar if the student's familiarity with "neuron," and related terms and concepts, is high). This can be referred to as "neighborhoods" of meaning within a semantic network.

Table 5 includes all 50 true terms, rank ordered as a function of the percent of hits. Instances of 100% performance on key terms across samples are potentially notable, as examples of concepts that all psychology majors can be expected to know. Several true terms were universally recognized: attachment, dependent variable, Jean Piaget, sexual identity, bell-curve, REM sleep, reinforcement, obsessive-compulsive disorder (OCD), and operant conditioning. If all the terms were labeled on the line graphs, these terms would appear at the far right end of the green curve in Figures 1 and 2. Of all of the foils, only one was completely unsuccessful at luring students: "hurtibility." This term would appear on the far left end of the red curve in Figures 1 and 2.

The questions posed to Auburn students regarding career interests produced the following results. Almost half circled "Other" as a possible career choice (n = 15). The next most popular choice was "Clinical" (n = 10), followed by "Academic/Teaching" (n

= 8) and “Industrial/Organizational” (*n* = 8). The least popular career choice was “Experimental” (*n* = 4). Table 6 presents additional responses to the career question.

DISCUSSION

The main hypothesis for this research, that performance on the APTT is a useful indicator of domain-specific learning, is supported by a) the relationships between graduating senior psychology major scores on the APTT and other measures of student performance (discussed further below), b) pre- and post-test differences in Introductory Psychology student scores (Smith, Brandt, & Barker, 2005), and c) the correlation between end of semester grades in Introductory Psychology and scores on the APTT, $r(257) = 0.63$, $p < .01$ (see Smith, 2006, p. 25). Figure 3 provides visual support for this assertion, by presenting APTT scores at points along the time course of acquiring a psychology degree (adapted from information reported in Barker, Smith, & Brandt, 2006, in press). It appears that terms presented in the Introductory Psychology course (and included in the APTT) are discriminated with more accuracy at final examination time (post-test) than at a midpoint (research methods extra credit opportunity mid-semester) between pre-test and graduating senior assessment (final post-test). The highest scores occur in students who have successfully completed a major in psychology.

Presumably, 4 years of coursework serves to increase familiarity with most of psychology's key terms. However, due to the nature of the degree requirements at Auburn, it is clear that the students may not have seen some of the terms since their first exposure in the Introductory class. The fact that the APTT scores do not fall from the levels originally obtained from in-class post-testing suggest that, insofar as this method of

assessing knowledge has been established as appropriate, hypothesis 7 (the APTT is an indicator of learning that persists over delay) has been supported. As well, hypothesis 2 (students from different institutions will not differ in their performance on the APTT) has also been supported.

It is clear that the APTT provides an approximate measure of the framework that education in the domain of psychology provides. It measures the degree of increase in discriminability of domain-specific key terms (and perhaps, in the confidence to interpret and use the language with certainty). It is unclear exactly how this maps onto their application of psychological concepts (and language) to everyday situations, or to narratives experienced via television screens, but the first attempt has been made to examine this question (hypothesis 7). The EVAψ scores did not correlate with other measures used to assess the students (ACT, study habits, in-class reading), but APTT scores (as well as GPA and intentions to go to graduate school) predicted greater quantity and quality of application of psychology terms to the psychological phenomenon in the cartoon. This provides important support for the assertion that the APTT is a highly useful assessment tool – performance on it correlates with performance on tasks that address three of the APA's guidelines for psychology major assessment.

The terms the students chose to write about are arguably terms that people who have never taken a course in psychology know, or if they do, they would not be capable of applying them accurately to *The Simpsons* (see Appendix C). Furthermore, d-prime helps explain differences among students grouped together based on ACT extremity – the highest ACT response s had only moderate d-prime scores (#1 response corresponds with d' = 1.88, and #2 with d' = 1.15, even lower than ACT lowest #1). The next two ACT

high responses, #3 and #4, which were scored as showing strong form-meaning connections, had APTT d' scores above the 85th percentile for the sample. Although ACT is considered a strong predictor of performance in college, it is certainly not a perfect predictor. In cases where it fails to identify the best performing students (on the written EVAψ), the APTT serves to provide a better approximation of psychology student knowledge.

Hypotheses 3 through 5 addressed whether ACT, GPA, and domain-specific reading contributed to APTT performance. Perhaps the strangest finding was that psychology GPA (arguably the most complete indicator of the consistency of performance in psychology classes) did not predict psychology key term discrimination. Because the rationale of calls for alternative methods of student assessment implies that GPA does *not* indicate enough about the status of knowledge and skills acquired through college education, this fact might provide the most support for the appropriateness of the design of the APTT. It does "tap" something different from what the earned grade indicates, at least for graduating seniors.

Since intelligence is sometimes conceptualized as "cognitive efficiency," and the process of verbal learning makes "great initial demands on conscious processing, but eventually become(s) automatic…(and) constitute(s) a fast-track neural pathway by which visual symbols can access the highest representational networks of the mind" (Donald, 2001, p. 304), it is not surprising that ACT was such a robust predictor of APTT performance. Particularly, higher ACT students were better able to reject the pseudo-terms, which suggests that they were better able to access entire "neighborhoods" of the semantic network and recognize that the pseudo-term did not appear in them.

The measure of cognitive efficiency provided by ACT scores provides a rough approximation of the degree to which students gain 'fast-track neural pathways' for psychology knowledge via coursework. The variables related to domain-specific reading (journal recognition and ASI-derived reading score, hypothesis 5) were not as successful at predicting APTT scores, but this must be due in part to deficits in the measurements' validity. The relationship between APTT d-prime and the d-prime for journals implies an overall confidence in student's awareness of things that are legitimately part of the field of psychology. However, the data suggest that the main sources of information at the undergraduate level remain professors, class lecture notes, and textbooks. Students can succeed at this level without awareness of the field's professional publications.

Hypothesis 6, the expectation that a meaning orientation to study was preferable to a reproduction orientation, was partially supported. No relationships between the meaning orientation and APTT performance were found. It is possible that the meaning-oriented students' focus on linking pieces of information, and unifying it all into a 'deeper meaning' framework, distracts them from knowing with certainty what is and what is not part of academic psychology.

As expected, students with a high reproduction orientation performed less well on the APTT, specifically because they were less able to correctly reject pseudo-terms. It should be noted that the items used to calculate the reproduction orientation are not necessarily indicative of the kind of item-specific processing that McCabe and colleagues describe, and that a combination of both techniques (processing for distinctive item knowledge and for relational system knowledge) is probably most effective for students engaged in the acquisition of a domain-specific semantic network (see Nation, 2001,

pp.297-302 and 320-343, for a discussion of the benefits and drawbacks of both decontextualized and 'contextualized' thought).

The limitations of this (mainly exploratory) study include a) the small number of testable hypotheses, b) the non-probability sampling procedure, c) the self-report nature of much of the data, d) the low response rate at Auburn, and e) the inability to control for or accurately measure the level of exposure to the terms (text reading, study time) and the time since exposure. Perhaps future research could incorporate measures of the same constructs, but that benefit (as the APTT yes-no recognition test methodology does) from immunity to social desirability effects.

Conclusion

As educators, we above all want students to become both senders and users of a wider range of concept representations. It can be argued that students arrive at college with many of psychology's concept representations already intact. However, they clearly do not possess a great deal of domain-specific vocabulary. Psycholinguists have posited that for every word or term, there exists both an underlying representation and a surface structure (the "chunk," or key term), and that for communication to be considered "successful," the underlying representation constructed by the receiver must match the underlying representation of the sender (Halpern, 1989, p. 89). Successful, efficient, and accurate thought and communication should be much more likely within a shared community of language users (i.e. people who have extensively studied psychology). The APTT assesses to what degree a student can be expected to utilize their domain-specific knowledge for sophisticated thought and successful communication after they leave the educational setting.

It is appropriate to claim that students who do well on the APTT have spent 4 years engaged in an intensive process of "filling in" their domain-specific semantic network. This results in a more-than-superficial awareness of the true versions of concept representations, such that they do not choose to accept the pseudo-terms. Even among students who do less well on the APTT, false alarms do not dominate their responding, as evidenced by the complete lack of negative d-prime values. This research provides support for the claim that graduates of psychology programs are more capable than non-psychology majors of discerning what is part of the finite domain of psychology terms from what is not. Furthermore, the APTT provides a tool to assess differences in the level of success that particular students, particular study habits, and/or particular teaching methods produce in discriminating psychology knowledge.

REFERENCES

Baumann, J.F., & Kameenui, E.J. (1991). Research on vocabulary instruction: Ode to Voltaire. In J. Flood, J.J.D. Lapp, & J.R. Squire (Eds.), *Handbook of research on teaching the English language arts* (pp. 604-632). New York: MacMillan.

Boneau, C.A. (1990). Psychological literacy: A first approximation. *American Psychologist, 45,* 891-900.

Coombs, C.H., Dawes, R.M., & Tversky, A. (1970). Signal detection. *Mathematical Psychology: An Elementary Introduction.* Englewood Cliffs, NJ: Prentice Hall.

Echols, L. D., West, R. F., Stanovich, K. E., & Zehr, K. S. (1996). Using literacy activities to predict growth in verbal cognitive skills: A longitudinal investigation. *Journal of Educational Psychology*, *88*, 296-304.

Entwistle, N.J., & Ramsden, P. (1983). *Understanding Student Learning.* London: Croon Helm.

Halonen, et al., (2002). The Assessment Cyberguide for Learning Goals and Outcomes in the Undergraduate Psychology Major. Retrieved September 14, 2004, from http://www.apa.org/ed/guide_outline.html

Halpern, D. F. (1989). *Thought and knowledge: An introduction to critical thinking.* Hillsdale, NJ: Lawrence Erlbaum.

Huibregtse, I., Admiraal, W., & Meara, P. (2002). Scores on a yes-no vocabulary test: correction for guessing and response style. *Language Testing, 19,* 227-245.

Lindsay, & Norman, (1977). *Human information processing: An introduction to psychology,* 2nd ed. New York: Academic Press.

MacMillan, N.A., & Creelman, C.D. (2005). *Detection Theory: A User's Guide,* 2nd ed. Mahwah: Lawrence Erlbaum.

McCabe, D.P., Presmanes, A.G., Robertson, C.L., & Smith, A.D. (2004). Item-specific processing reduces false memories. *Psychonomic Bulletin and Review, 11*, 1074-1079.

Meara, P., & Fitzpatrick, T. (2000). Lex30: An improved method of assessing productive vocabulary in an L2. *System, 28,* 9-30.

Merlin, D. (2001). *A mind so rare: The evolution of human consciousness*. New York: Norton.

Mochida, A. & Harrington, M. (2006). The yes/no test as a measure of receptive vocabulary knowledge. *Language Testing, 23,* 73-98.

Nation, I.S.P. (2001). *Learning vocabulary in another language.* Cambridge: Cambridge Univesity Press.

Newstead, S. (1992). A study of two "quick-and-easy" methods of assessing individual differences in student learning. *British Journal of Educational Psychology, 62,* 299-312.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology, 2004,* 87-124.

Siddiqui, S., West, R.F., & Stanovich, K.E. (1998). The influence of print exposure on syllogistic reasoning and knowledge of mental state verbs. *Scientific Studies of*

Reading, 2, 81-96.

Stanovich, K. E. & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology*, *85*, 211-229.

Smith, D.L. (2006). *Development and validation of the Auburn Psychology Term Test (APTT).* Unpublished master's thesis, Auburn University, Auburn.

Smith, D.L., Brandt, J.A., & Barker, L. (2005, February). *A recognition test for assessment of psychology students*. Poster session presented at the annual meeting for Southeastern Teaching of Psychology, Atlanta, GA.

Tanner, W. P. & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review, 61,* 401-409.

West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, *28*, 34-50.

White, K.G. & Wixted, J.T. (1999). Psychophysics of remembering. *Journal of the Experimental Analysis of Behavior, 71,* 91-113.

Wixted, J.T. & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin and Review, 11,* 616-641.

Appendix A: Auburn Psychology Term Test (Alpha)

Part One – (First, please bubble in your codename on the provided scantron)

Your task is to identify which of the terms you recognize as legitimate psychology terms. Bubble in "A" if you recognize it as a real term, and "B" if you think the term is bogus:

1. negative reward
2. chunking
3. antisocial facilitation
4. attachment
5. aversive stimulus
6. law of effort
7. threshold of non-relativity
8. dependent variable
9. Patricia Broca
10. transduction
11. absolute threshold
12. unconditional negative regard
13. general intelligence (*g*)
14. cognitive dissonance
15. critical period
16. cropping response
17. dark adaptation
18. dendrite
19. token economies
20. transperceptible difference (TD)
21. Gestation psychology
22. standard deviation
23. Jean Piaget
24. dendritic hypo-potential
25. negative feedback
26. sexual identity
27. libido
28. superstitious relaxation
29. bell-curve
30. sensory chunking
31. neurostasis
32. fixation
33. hidden observer
34. hapless motivation
35. schema taking score (STS)
36. multiple deviation
37. involutional study
38. REM Sleep
39. shaping
40. Sigmund Froyd
41. circadian movement
42. proto-operational stage
43. James-Lange theory
44. retrograde memory
45. Wernicke's area
46. longitudinal study
47. somatic transmission
48. psychotransference
49. reinforcement
50. neuroresponder
51. instinctual deprivation
52. indifferent schizophrenia
53. unconscious neurotocism
54. successive approximation
55. imprinted neurosis
56. obsessive compulsive disorder (OCD)
57. operant conditioning
58. proactive interference
59. terminal stasis
60. distance IQ
61. John B. Watson
62. interdependent variable
63. strange situation test
64. hurtibility
65. law of effect
66. RPM effect
67. magneto-convulsive therapy
68. unsystematic sensitization
69. operational definition
70. learned affiliation
71. long term memory
72. circadian rhythm
73. clunking potential (CP)
74. spontaneous salivation
75. fundamental attribution error
76. opponent-process theory
77. Festinger-Maslow Effect
78. just noticeable difference (JND)
79. William James
80. echoic inhibition
81. sensorimotor stage
82. transferential psychosurgery
83. introspection
84. duozygotic twins
85. Elizabeth Loftus
86. ego complex
87. specific adaptation syndrome (SAS)
88. episodic memory
89. test-retest reliability
90. unconditioned restriction
91. operant encoding
92. systematic desensitization
93. psychophysics
94. post-distress disorder
95. latent gratification
96. fixed action pattern (FAP)
97. social loafing
98. captive nerve ending
99. hypostasis
100. action potential

Appendix B: Auburn Psychology Term Test (APTT) key, with true terms bolded

1. negative reward
2. chunking
3. antisocial facilitation
4. attachment
5. aversive stimulus
6. law of effort
7. threshold of non-relativity
8. dependent variable
9. Patricia Broca
10. transduction
11. absolute threshold
12. unconditional negative regard
13.general intelligence (*g*)
14. cognitive dissonance
15. critical period
16. cropping response
17. dark adaptation
18. dendrite
19. token economies
20. transperceptible difference (TD)
21. Gestation psychology
22. standard deviation
23. Jean Piaget
34. hapless motivation
35. schema taking score (STS)
36. multiple deviation
37. involutional study
38. REM Sleep
39. shaping
40. Sigmund Froyd
41. circadian movement
42. proto-operational stage
43. James-Lange theory
44. retrograde memory
45. Wernicke's area
46. longitudinal study
47. somatic transmission
48. psychotransference
49. reinforcement
50. neuroresponder
51. instinctual deprivation
52.indifferent schizophrenia
53. unconscious neurotocism
54. successive approximation
55. imprinted neurosis
56. obsessive compulsive disorder
68. unsystematic sensitization
69. operational definition
70. learned affiliation
71. long term memory
72. circadian rhythm
73. clunking potential (CP)
74. spontaneous salivation
75. fundamental attribution error
76. opponent-process theory
77. Festinger-Maslow Effect
78. just noticeable difference (JND)
79. William James
80. echoic inhibition
81. sensorimotor stage
82. transferential psychosurgery
83. introspection
84. duozygotic twins
85. Elizabeth Loftus
86. ego complex
87. specific adaptation syndrome (SAS)
88. episodic memory
89. test-retest reliability
90. unconditioned restriction

(OCD)

24. dendritic hypo-potential

25. negative feedback
26. sexual identity
27. libido
28. superstitious relaxation

29. bell-curve
30. sensory chunking
31. neurostasis
32. fixation
33. hidden observer

57. operant conditioning
58. proactive interference
59. terminal stasis
60. distance IQ

61. John B. Watson
62. interdependent variable
63. strange situation test
64. hurtibility
65. law of effect
66. RPM effect
67. magneto-convulsive therapy

91. operant encoding
92. systematic desensitization
93. psychophysics
94. post-distress disorder

95. latent gratification
96. fixed action pattern (FAP)
97. social loafing
98. captive nerve ending
99. hypostasis
100. action potential

Appendix C: Approaches to Studying Inventory

Part Two – Please answer the following questions, mostly about your study habits over the past several years, and your future academic plans and interests, honestly and objectively.

Circle a number to indicate the extent to which you agree or disagree with these statements: Disagree (1) Neutral (3) Agree (5)

1. I find it easy to organize my study time effectively.
2. I like to be told precisely what to do in essays or other set work.
3. It's important to me to do really well in the courses here.
4. I usually set out to understand thoroughly the meaning of what I am asked to read.
5. When I'm reading I try to memorize important facts which may come in useful later.
6. When I'm doing a piece of work, I try to bear in mind exactly what that particular lecturer seems to want.
7. My main reason for being here is so that I can learn more about the subjects which really interest me.
8. I suppose I'm more interested in the qualifications I'll get than in the courses I'm taking.
9. I'm usually prompt in starting work in the evenings.
10. I generally put a lot of effort into trying to understand things which initially seem difficult.
11. Often I find I have to read things without having a chance to really understand them.
12. If conditions aren't right for me to study, I generally manage to do something to change them.
13. I often find myself questioning things that I hear in lessons or read in books.
14. I tend to read very little beyond what's required for completing assignments.
15. It is important to me to do things better than my friends, if I possibly can.
16. I spend a good deal of my spare time in finding out more about interesting topics which have been discussed in class
17. I find academic topics so interesting, I should like to continue with them after I finish this course.
18. I find I have to concentrate on memorizing a good deal of what we have to learn.

What was your overall GPA (estimate is okay)?_____

What was your GPA in psychology courses (estimate is okay)?_____

What was your ACT score (estimate is okay)?_____

Do you plan to go to graduate school in psychology (circle one)? Yes No Maybe

Appendix D-1: Writing Samples from highest five APTT d' scores (2.12 to 3.11)

d' = 3.11

- the law of effect – BARELY related, but in killing species in the past, different consequences/results
- conditioning – use "brainwashing" to get everyone to be like Flanders. punished for wrong responses.
- frontal lobotomy – in through the nose? Not really. Also no brain cut out, usually use a tool to destroy tissue – go in through ocular cavity
- evolution – fish crawls out of water, kinda funny
- Homer's denial and acceptance, stages of grief?

d' = 2.97

Our perception of the world is what we make of it. Everyone has their own idea of what is correct or normal for them. As Homer continues to change things in the past his future is changed. In his view sometimes for the better. Ultametaly he finds he accepts what was his "normal" life. I'm sure there are many concepts throughout the episode but I could not really think of any.

d' = 2.49

The cartoon clip included the phenomenon of time travel and portrayed Homer traveling back to the time of the dinosaurs using a broken toaster. This is not necessarily a psychological concept (it's more of a scientific one) but the way Homer thinks about the consequences of time travel is related. He remembers the advice his father gave him about not touching or changing anything in the past and believes his father is right. In one return to "the present," Ned Flanders has taken over and is Re-Neducating everyone in the Flanders way by removing everyone's frontal lobe in a frontal lobotomy, which is an actual neuro-surgical procedure. However, it does not necessarily result in apathy, forgetfulness, or lethargy as was depicted in the members of the Simpson family.

d' = 2.32

Memory – Advice from his father
Peer pressure – Bart influencing Lisa to kill Homer
Learning – Relational Abstract Concept of toaster and time travel
Reinforcement – Homer's reinforcement history influences his decision of whether to use the toaster to travel through time again.

d' = 2.12

Time travel and all its crazy effects are explored. He goes through all different worlds with different aspects of Homer's life that have been altered. Also, hypnotizing is covered when Flanders has hypnotized the whole town and done labotomies on them to control their thoughts. I don't really think that is a realistic view of lobotomy or hypnotizing. Labotomy is not done to control peoples minds and it is hardly ever really done anymore. Hypnotizing is not nearly as dramatic as it is depicted in this clip. It is used mildly to help people with issues such as smoking. I also think it was interesting when his children were huge and called him a bug. Maybe each world he entered was a different part of his unconscious.

Appendix D-2: Writing Samples from lowest five APTT d' scores (0.66 to 0.94)

d' = 0.66
(Note: this person wrote 3 letters – hyp – then erased them, leaving a blank page)

d' = 0.71
schizophrenia
cognitive
Adaptation

d' = 0.74
- might be having deliusions/hallucinations
- positive/negative feedback
- mentioning of different areas of the brain
- flashback episoide
- operant conditioning
- Death – Grieving period (Developmental Psyc.)
- psychotic behavior
- social interaction w/in the family
- type of family

d' = 0.84
Homer kept messing with dependent variables when he went back in the future, by killing the fish it changed the future & his life. There were many noticeable differences that occured and at the very end everything seemed normal except for one just noticeable difference of their tongue & how they ate. It was almost like he was hallucinating in some of the scences which w/out drugs are symptoms of disorder.

d' = 0.94
Mental retardation
Halucinations
OCD
Frontal lobatamy? – hopefully doesn't happen
obesity

concept:
everything you do effects other things and people

Appendix D-3: Writing samples grouped according to non-APTT characteristics

Grade Point Average (4.0) Responses:

1. The cartoon clip included the phenomenon of time travel and portrayed Homer traveling back to the time of the dinosaurs using a broken toaster. This is not necessarily a psychological concept (it's more of a scientific one) but the way Homer thinks about the consequences of time travel is related. He remembers the advice his father gave him about not touching or changing anything in the past and believes his father is right. In one return to "the present," Ned Flanders has taken over and is Re-Neducating everyone in the Flanders way by removing everyone's frontal lobe in a frontal lobotomy, which is an actual neuro-surgical procedure. However, it does not necessarily result in apathy, forgetfulness, or lethargy as was depicted in the members of the Simpson family.

2. persuasion/conformity – as "slaves" of Ned Flanders
   personal fable – Homer thinks that what happens to him has never happened to anyone and that experiences are unique.
   attachment styles of the family
   aggression – when he clubs everything, yells at Marge
   stereotype – the aliens assume all humans are unprepared for effects of time travel

3. Most obviously misconstrued in this cartoon is the idea that our seemingly inconsequential actions can change the future drastically. While this can sometimes be true in some sense, we cannot base our decisions on this notion without inducing severe stress. Homer got extremely upset about silly things, but became so exhausted that he didn't even care that his family had lizard tongues in the end. This is a case of learned helplessness.

Grade Point Average (range 2.5 - 2.8) Responses:

1. - frontal lobe
   - people's susceptibility to following the group; such as in riots.
   - schizophrenia – in own world
   - manic – very high and energized, very excitable

2. The cartoon could represent socialization. What we are socialized to believe is the "norm" in society. Each time Homer comes back his family is different then what he is used too. He doesn't adapt to the new changes if he cannot handle them. It could also portray stereotypes of men & women. Marge is in the kitchen doing the cooking. Homer gets his hand stuck in the toaster, possibly portraying that men should not or are not suppose to be in the kitchen.

3. (Note: this person wrote 3 letters – hyp – then erased them, leaving a blank page)

ACT Highest* (28 – 31) Responses

*The GPA highest responses include 2 additional high ACT scorers:
#3 reported an ACT of 35, and #1 reported an ACT of 30

1. Lobotomy? Wow, I could go on about that one… Also, the idea possibly of psychology as "reprogramming"? Or am I just bitter?

2. BELIEF PERSEVERANCE
   FRONTAL LOBOTOMY
   SOCIAL LEARNING THEORY

3. - the law of effect – BARELY related, but in killing species in the past, different consequences/results
   - conditioning – use "brainwashing" to get everyone to be like Flanders. punished for wrong responses.
   - frontal lobotomy – in through the nose? Not really. Also no brain cut out, usually use a tool to destroy tissue – go in through ocular cavity
   - evolution – fish crawls out of water, kinda funny
   - Homer's denial and acceptance, stages of grief?

4. Memory – Advice from his father
   Peer pressure – Bart influencing Lisa to kill Homer
   Learning – Relational Abstract Concept of toaster and time travel
   Reinforcement – Homer's reinforcement history influences his decision of whether to use the toaster to travel through time again.

ACT Lowest (19)

1. Compliance
   Conformity

2. schizophrinia
   cognitive
   Adaptation

Meaning Orientation Highest Score* (20-21) Responses

*Several high scorers' responses appear elsewhere: the highest Meaning scores (24 and 22) correspond to responses reported as ACT highest #1, #2, and GPA highest #1. Another high Meaning score (21) corresponds to ACT lowest #2. Two more high Meaning scores (20) already appear as ACT highest #4 and GPA lowest #3.

1. Mental retardation
   Halucinations
   OCD
   Frontal lobatamy? – hopefully doesn't happen
   obesity

concept: everything you do effects other things and people

2. – Memory
   - Programming?
   - No such thing as a frontal lobotomy.
   - The current lobotomy techniques don't go through the nose.
   - Aggression

3. The idea that positive punishment makes a action less likely is misrepresented. Simpson keeps touching the toaster. A labodomy makes people less irritable.

Meaning Orientation Lowest Score (10-14) Responses

1. <u>cognitive dissonance</u> → people are shaping beliefs and their lives to conform to what is acceptable to society. Homer, on the other hand is not.
   He becomes obsessive compulsive with the toaster to escape reality & go back in time.

2. One psychological concept that I see illustrated is that of Homer not being in control of his life. Everything around him keeps changing, and he cannot adapt to it. He wants his life and his family to go back to the way they were before he went back in time. The cartoon shows Homer feeling crazy and annoyed that things have changed around him.

Reproduction Orientation Highest* (21-23) Score Responses

*One highest (23) score corresponds with GPA lowest response #2. Another high score (22) corresponds with ACT lowest #2

1. - Homer recalled what his father told him on his wedding day.
   - He began to be driven crazy because of unfamiliar surroundings.
   - At the beginning, the family was having supper together → positive interaction.
   - At one point, everyone was lined up listening to some man on television as if they had been brainwashed & were not thinking for themselves. Homer decided to get away from the group which could illustrate his noncompliance.
   - Social learning / modeling could occur for the children, because they see how Homer acts then act it out themselves.

2. - Negative Punishment – saying something bad about the guy in control results in punishment.
   - Homer is acting in a paranoid nature. Its like he has a lot going on in his brain, (schziophrenia?)

3. - might be having deliusions/hallicinations
   - positive/negative feedback
   - mentioning of different areas of the brain
   - flashback episoide
   - operant conditioning
   - Death – Grieving period (Developmental Psyc.)
   - psychotic behavior
   - social interaction w/in the family
   - type of family

Reproduction Orientation Lowest* (14) Score

*The lowest score (11) corresponds with GPA highest #3. ACT Highest #1, Lowest #1 and Meaning Highest #3 also had low (12) scores. Another low score (14) corresponds with ACT highest #4.

1. - making them smile physically to produce elevated mood
   - performing a lobotomy to gain control over them

2. Homer killing even one small bug causes the whole future to change. This tells about how some people believe if they do one small thing, it can change their lives forever. Some believe the world is linked together and every action produces a reaction.

Reading High Score* (3) Responses

*Reproduction lowest #2 and GPA highest #3 correspond to the highest reading score (5). The response for ACT highest #4 also scored high (4). Five responses appearing elsewhere – Meaning high #1 and #3, ACT low #2 and high #1, and GPA lowest #3 – correspond with a Reading score of 3.

1. The Simpson's humor has several subliminal messages in it about all kinds of subjects. In this episode, they construe how the general population/public views full frontal labotomies. When Flander's puts the family in the theater & forces them to smile, this reminds me of how anti-depressants tend to work/effect some people (fake smile – fake happiness). When Homer hears his father in his head can represent subconscious thoughts. Overall, Homer viewing the world/his family in different/strange ways can somewhat represent hallucinations or even go as far as schizophrenia & paranoia.

2. Time travel and all its crazy effects are explored. He goes through all different worlds with different aspects of Homer's life that have been altered. Also, hypnotizing is covered when Flanders has hypnotized the whole town and done labotomies on them to control their thoughts. I don't really think that is a realistic view of lobotomy or hypnotizing. Labotomy is not done to control peoples minds and it is hardly ever really done anymore. Hypnotizing is not nearly as dramatic as it is depicted in this clip. It is used mildly to help people with issues such as smoking.
   I also think it was interesting when his children were huge and called him a bug. Maybe each world he entered was a different part of his unconscious.

Reading Low Score* (-3 and -1) Responses

*The following appear elsewhere: GPA low #2 (-5); GPA high #1 (-4); Meaning low #1 (-3); Reproduction high #1, ACT low #2, GPA low #1 (-1)

1. Homer kept messing with dependent variables when he went back in the future, by killing the fish it changed the future & his life. There were many noticeable differences that occured and at the very end everything seemed normal except for one just noticeable difference of their tongue & how they ate. It was almost like he was hallucinating in some of the scences which w/out drugs are symptoms of disorder.

2. Unethical experimenting on Simpsons.
   Fundamental attribution error.

Responses Not Appearing Elsewhere:

1. Time travel? Possibility
    1. Homer = Pain sense, his imagination, eating gives energy
    2. Ned = Power Hunger
    3. Brain Parts & functions
    4. Actions change everything w/ time
    5. Maggie – Violent w/ axe
    6. Unnatural Unrealistic family life
    7. evolutionary process
    8. Behavioral Neuroscience

    This show basically portrays how creative the writers can be providing that it is all unrealistic & imaginative. Removing brain parts alters the behavior & actions of people. Ned shows a very power hungry attitude in somewhat of a Hilter manner. This clip demonstrates how different actions can affect the evolutionary process as a whole. A small child, Maggie, should not be allowed to play w/ dangerous objects (ex. axe) & kill people. Homer has a wild incredible imagination & in the beginning his whole sense of pain & their receptors seem to be out of whack, but then he shows how the man of the house always tries to fix things & sometimes just makes them worse. All in all at least he tried and did not give up on his main goal.

2. Theory of Evolution – Homer going back in time and messing up the future by killing the mosquito. Ned is found to be the ruler of the world.

   Learning & Reinforcement – Homer figuring out that if he touched or destroyed anything the future would be different. He had to go through several trials to figure out how to get back to the old world (normal Springfield) but I don't think he ever figured it out.

3. Conformity → when the guy who controls everything
   Roles → they all have a role in the family and part.

4. Our perception of the world is what we make of it. Everyone has their own idea of what is correct or normal for them. As Homer continues to change things in the past his future is changed. In his view sometimes for the better. Ultametaly he finds he accepts what was his "normal" life. I'm sure there are many concepts throughout the episode but I could not really think of any.

Table 1: Percent recognition for Auburn students on the journal recognition task

Journal Title	Percent Recognition
Behavioral Neuroscience	72
Developmental Psychology	63
Health Psychology	75
History of Psychology	25
Journal of Educational Psychology	50
Journal of Experimental Psychology	59
Journal of Personality and Social Psychology	41
Neuropsychology	38
Psychological Bulletin	31
Psychological Review	78

Journal Foil Title	Percent Correct Rejection
Abnormal Psychology Journal	44
America's Psychologist	63
Cultural Diversification and Multi-Minority Psychology	100
Emotional Psychology	88
Journal of the Application of Psychology	72
National Journal of Stress Management	94
Psychoanalytical Integration Psychology	100
Psychology and Law	78
Psychology and Lifespan Development	69
Systems, Families and Health	88

Table 2: Descriptive statistics for Auburn and Elizabethtown students on the APTT, ASI, and other measures

	Auburn M	SD	Elizabethtown M	SD
1. Hits	40.94	4.46	39.36	6.65
2. Correct Rejections	35.55	8.03	40.55	6.61
3. d'	1.60	0.62	1.91	0.75
4. GPA psychology	3.42	0.41	3.29	0.53
5. GPA total	3.23	0.38	3.35	0.53
6. ACT	24.72	4.10	28.33	3.88
7. Reading	0.90	2.48	1.82	2.14
8. Meaning	18.19	2.80	18.82	2.52
9. Reproduction	17.10	3.33	12.73	2.83

Table 3: Pearson 2-tailed correlations among variables of interest for the combined sample

	M	SD	1	2	3	4	5	6	7	8	9
1. Hits	40.83	5.0	X	-.12	.19	.28	.16	.14	.16	-.18	-.16
2. Correct Rejections	36.55	7.8		X	.54*	.16	.28	.51*	.11	.26	-.38
3. d'	1.68	0.7			X	.28	.34*	.54*	.11	.07	-.38*
4. GPA psychology	3.39	0.4				X	.79*	.45	.35	.23	-.29
5. GPA total	3.26	0.4					X	.39	.31	.20	-.33
6. ACT	25.1	4.2						X	.20	.41	-.35
7. Reading	1.14	2.4							X	.31	-.63
8. Meaning	18.4	2.7								X	-.27
9. Reproduction	16.83	3.4									X

Notes: Underlined values indicate $p < .05$ (2-tailed significance)
* indicates $p < .01$

The *n* for these correlations varies as follows: ACT = 26 (only 26 of 31 in the Auburn sample reported it, and only 3 of the 11 in the Elizabethtown sample reported scores [SAT]. Both correlations [with *n* of 26 and 29] were computed, with generally higher values but similar patterns found when the 3 additional scores were added. In the interest of presenting a more accurate representation of the relationships, only the Auburn sample's scores are used for these and other computations. For all other correlations, the sample size is 42.

Table 4: Foils rank-ordered by percentage correct responding for combined sample

hurtibility	100%	retrograde memory	24%
transferential psychosurgery	97%	negative reward	32%
clunking potential (CP)	95%	ego complex	36%
involutional study	95%	latent gratification	36%
hapless motivation	94%	learned affiliation	40%
cropping response	93%	spontaneous salivation	50%
transperceptible difference (TD)	90%	somatic transmission	57%
superstitious relaxation	90%	multiple deviation	57%
unsystematic sensitization	90%	psychotransference	57%
law of effort	88%	sensory chunking	60%
magneto-convulsive therapy	88%	operant encoding	64%
instinctual deprivation	88%	circadian movement	64%
threshold of non-relativity	86%	interdependent variable	67%
schema taking score (STS)	86%	Patricia Broca	67%
distance IQ	86%	captive nerve ending	69%
echoic inhibition	86%	neurostasis	69%
unconditional restriction	83%	dendritic hypo-potential	69%
duozygotic twins	83%	antisocial facilitation	69%
RPM effect	83%	Sigmund Froyd	71%
imprinted neurosis	83%	unconscious neurotocism	71%
neuroresponder	83%	post-distress disorder	74%
proto-operational stage	81%	Festinger-Maslow effect	74%
hypostasis	79%	terminal stasis	74%
unconditional negative regard	79%	indifferent schizophrenia	76%
Gestation psychology	76%	specific adaptation syndrome (SAS)	76%

Table 5: True terms rank-ordered by percent correct for combined sample

strange situation test	38%	attachment	100%
dark adaptation	40%	dependent variable	100%
fixed action pattern (FAP)	43%	sexual identity	100%
Elizabeth Loftus	43%	bell-curve	100%
successive approximation	46%	REM sleep	100%
psychophysics	57%	reinforcement	100%
James-Lange theory	57%	obsessive-compulsive disorder (OCD)	100%
hidden observer	60%	operant conditioning	100%
proactive interference	60%	long term memory	100%
just noticeable difference (JND)	62%	Jean Piaget	100%
transduction	63%	longitudinal study	98%
chunking	64%	standard deviation	98%
circadian rhythm	69%	negative feedback	98%
law of effect	76%	fixation	95%
William James	76%	cognitive dissonance	95%
social loafing	76%	critical period	95%
token economies	78%	shaping	95%
opponent process theory	79%	introspection	93%
Wernicke's area	79%	libido	93%
fundamental attribution error	81%	test-retest reliability	91%
systematic desensitization	83%	action potential	91%
dendrite	83%	aversive stimulus	90%
absolute threshold	83%	operational definition	88%
episodic memory	86%	general intelligence (g)	88%
sensorimotor stage	86%	John B. Watson	86%

Table 6: Additional Responses to Career Interests Question

Psychology-related	Unrelated
Infant cognition lab at U.T. Austin	real estate
Developmental disabilities	business relations
Field research	law/military
Professional counseling	nurse practitioner
Neuroendocrinology/psychopharmacology	elementary education
Clinical work with juvenile delinquents	medical doctor
Human development and family studies/prevention services	teaching Spanish
Autistic early childhood	teaching math
Christian counseling; marriage and family therapy	
Like abnormal; like working directly with children	
Adolescent therapy	
Teaching in high school	
Private practice	
Forensic psychology	
School counseling	

Figure 1: Probability of graduating seniors recognizing psychology terms (green curve) and recognizing pseudo-terms (red curve), plotted as a function of familiarity.

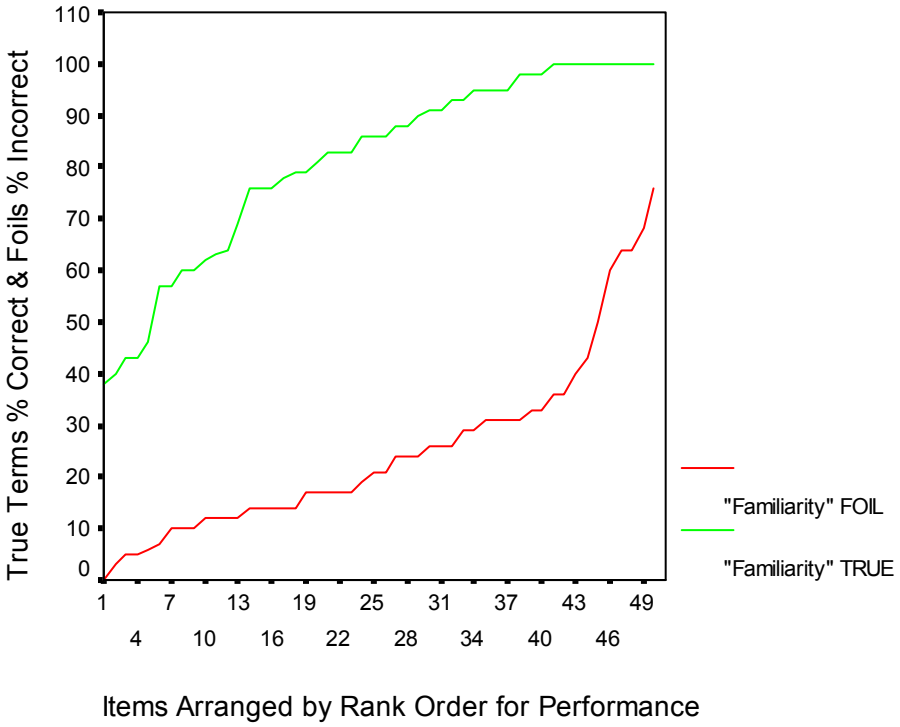

Figure 2: Probability of recognition of psychological terminology as a function of familiarity. The terms represented in the middle region (see Tables 4 and 5) correspond to the overlapping regions of the S and S+N distributions in signal detection theory.

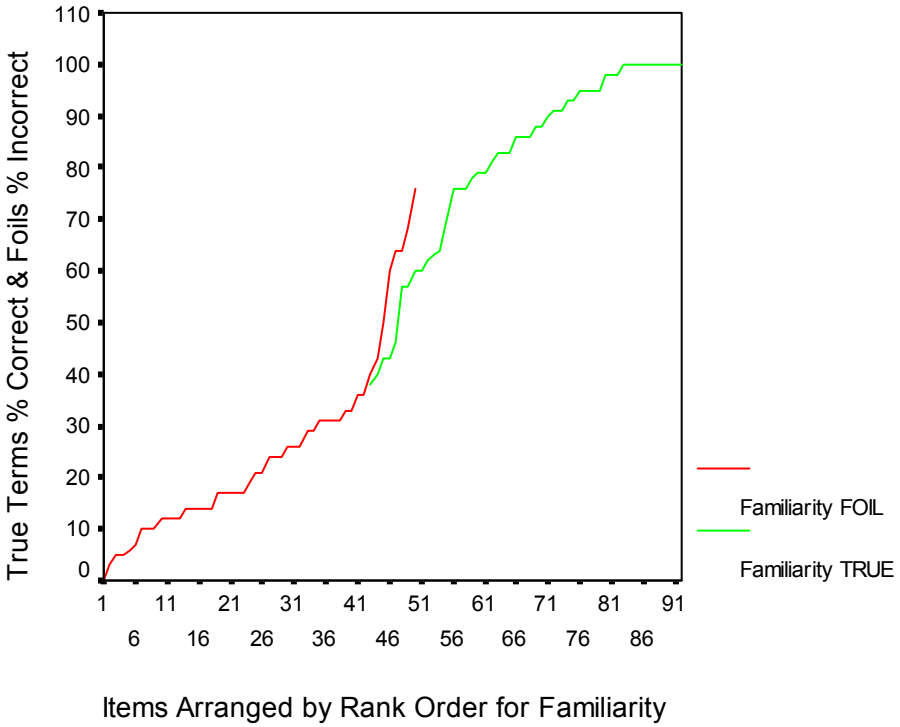

Figure 3: Raw scores on the APTT, at different points in the time course of obtaining the psychology degree