2008

Fifty Years of Memory of College Grades: Accuracy and Distortions

Harry P. Bahrick
Ohio Wesleyan University, hpbarick@owu.edu

Lynda K. Hall
Franklin University, lynda.hall@franklin.edu

Laura A. Da Costa
Fresno City College, ladc92@yahoo.com

This is a prepublication author manuscript. The published article is available at: https://doi.org/10.1037/1528-3542.8.1.13

Running Head: MEMORY OF COLLEGE GRADES

Fifty Years of Memory of College Grades: Accuracy and Distortions

Harry P. Bahrick, Lynda K. Hall

Ohio Wesleyan University

and

Laura A. Da Costa

Abstract

One to 54 years after graduating, 276 alumni correctly recalled 3025 of 3967 college grades. Omission errors increased with the retention interval, and better students made fewer errors. Accuracy of recall increased with confidence in recall. Eighty one percent of commission errors inflated the actual grade. Distortions occur soon after graduation, remain constant during the retention interval, and are greater for better students and for courses students enjoyed most. Confidence in recall is unrelated to distortion. Courses that were not freely recalled, but had to be cued, were recalled less accurately and with less distortion. The data support a supplementary theory of memory distortion. The theory assumes that forgetting and distorting memory content are relatively independent processes; that relevant generic memories are used to fill in gaps after episodic memory fails; that systematic distortions affect autobiographical content that is emotionally and motivationally valenced, and that most individuals supplement with content that is emotionally more gratifying than the veridical content. The data conflict with dynamic displacement theories according to which screen memories actively block access to unpleasant veridical content.

Fifty Years of Memory of College Grades: Accuracy and Distortions

In this report, we describe an investigation that focuses on both accuracy and distortions of memory. James recognized long ago (1890, p.373) that autobiographical memories are frequently distorted if there is a strong motivational and emotional investment in their content. However, in spite of their acknowledged significance, emotionally driven, self generated distortions of memory are not often investigated. The past neglect is due to methodological challenges (Bahrick, 1996, 1997; Koriat and Goldsmith, 1994).

The topic of false memories and memory distortions has received much recent attention (e.g., Roediger & McDermott, 1995). The impetus for the recent surge of false memory research stems from litigation and public controversy regarding the accuracy of recovered childhood memories of alleged sexual abuse. Such experiences are likely to occur in an emotionally charged context, and relevant research on the dynamics of distortion needs to explore the relationship of the emotions and motivations generated by the original experience to subsequent distortions of content. More specifically, investigators need to address the basis for the direction, magnitude, and temporal course of distortions. Experiences involving strong motivation/emotion are difficult to generate in the laboratory, and therefore this type of research does not lend itself readily to the experimental paradigm. Emotionally charged experiences occur primarily in naturalistic settings where they are likely to be unique to the individual and difficult to monitor, verify or replicate. These characteristics present formidable methodological obstacles.

Koriat and Goldsmith (1994) contrasted the storehouse and correspondence metaphors of memory. They pointed out that the storehouse metaphor is appropriate for conceptualizing quantitative losses of content that consists of individual units, e.g., nonsense syllables or words. This approach is illustrated by traditional memory research beginning with Ebbinghaus. The correspondence metaphor is appropriate to conceptualize qualitative changes typically involved in the reconstruction of more complex autobiographical content that does not necessarily consist of comparable, individual units. Bartlett's (1932) investigation of systematic changes in the recollection of stories told to children of various cultures is an example. Such memory distortions are more difficult to quantify. If emotionally driven systematic changes are involved, relevant theory must address the dynamics of change.

In an excellent demonstration of motivated memory distortions, Neisser (1981) examined systematic differences between two versions of the content of conversations between John Dean and President Nixon. One version was the taped record kept in the White House; the other was recalled by John Dean in his testimony at the

congressional Watergate hearings. Neisser's analysis documents how Dean's motives affected the nature of his distortions, but the complex content of the conversations and the limitation to the memory of a single individual make it difficult to quantify or generalize the results.

In a previous investigation (Bahrick, Hall & Berger, 1996), we were able to quantify systematic, emotionally driven changes in autobiographical memory by examining the recall of 3220 high school grades from 99 college students. We found that the participants made 814 errors of commission and that 82% of their errors were in an inflationary direction; i.e., the recalled grade was higher than the grade recorded on the high school transcript. There are several reasons why grades provide a suitable content for exploring the dynamics of systematic memory distortions. Grades are often encoded in an emotional context of satisfaction or frustration, depending upon the relation of the grade to the individual's expectations, aspirations, etc. Also, grades are quantifiable and verifiable and comparably scaled among individuals. A further important advantage is that the use of grades avoids the problematic confounding of memory distortions with perceptual distortions. It is highly likely that grades are originally encoded correctly, because it is highly unlikely that individuals perceive a grade different from the one shown when they first see their grade report. Emotional autobiographical experiences may be distorted because of misperceptions at the time of encoding. It is therefore an important methodological challenge for memory scholars to be able to differentiate such distortions of perception from distortions of memory that occur during the retention interval.

The goals of the present study were to extend the earlier investigation (Bahrick et al., 1996) by examining the relation of memory distortions of grades to several additional variables and at the same time to test predictions derived from the 'supplementary theory' of distortions proposed in the earlier investigation. The Bahrick et al. (1996) investigation compared two alternative explanations of systematic distortions. The supplementing theory (Bahrick, 1997) asserts that recall of autobiographical memory involves both replicative (reproductive) and reconstructive processing (Bahrick, 1984a, 1984b) and that reconstructive processing is used to fill in the gaps when attempts at replicative processing fail. The theory assumes that distortions occur during reconstructive processing, that they are based upon relevant generic knowledge, and that the distorted content is likely to generate more positive affect than the veridical content. Systematic modifications are most likely to involve content that is negatively valenced because such content is less likely to be rehearsed and therefore has a shorter life span and because such content lends itself more readily to changes that will generate more positive affect than the veridical

content. As an example, a student may no longer be able to recall the grade she received in the second semester of a course in geometry, but the generic knowledge that she was a good math student leads her to reconstruct a grade of A more frequently than a grade of C when the actual grade was a B.

This view of distortions contrasts with a dynamic replacement theory of distortion, according to which reconstructive processing actively interferes with or blocks access to the veridical memory content. The replacement view is in accord with Gestalt theory and with psychoanalytic theory. Gestalt theorists (Wulf, 1922; Bartlett, 1932) assumed that sharpening, leveling and assimilating are tendencies that systematically distort memory content so as to bring the content into better accord with existing forms or schemata. Psychoanalytic theory stipulates that distortions affect primarily traumatic or emotionally disturbing content. Screen memories actively displace anxiety producing content (Freud,1899/1962). As Edwards (1942) described it, "Events are forcefully obliterated from consciousness because they are not in harmony with the ego's ideals, wishes and values" (p.43). The supplementary and replacement theories may not be mutually exclusive. Dynamic displacement of content may occur primarily in connection with traumatic experiences, while supplementary modifications may predominate for content that generates less intense emotion.

The Bahrick et al. data showed that errors of recall are much less frequent for high grades than for low grades and that the proportion of inflationary errors is higher for better students. We attributed the finding that higher grades are remembered more accurately than lower grades to more frequent rehearsals that generate positive affect and prolong the life-span of gratifying content. However, recent findings of Guy and Cahill (1999) and Hamman (2001) indicate that more frequent rehearsals are insufficient to account for superior memory of emotional events and that the amygdala is selectively involved in enhancing such memory content.

The finding that the degree of distortion in the positive direction is greater for better students seems counterintuitive. We attributed it to generic memories. Better students have longer lasting memories of their emotionally satisfying grades and therefore fewer occasions to reconstruct them. However, once the episodic memory of a specific grade is lost, reconstructions based on the generic memories of many high grades yield a higher probability of positive distortions than the generic memories of poorer students.

The assumption that memory reconstructions are motivated to enhance positive affect (Loftus, 1982) leads to further predictions examined in the present investigation. First, we predicted that the degree of distortion should reflect the intensity of emotion associated with the memory content. We assumed that students are motivationally

more invested in courses that pertain to their major than in other courses, and we compare distortion of grades for the two types of courses. The present study also extends the earlier findings by examining accuracy and distortion of grades over a 54 year retention interval. These data offer the opportunity to examine the time-course of distortions during this long period. Because the motivational investment and emotional intensity associated with receiving grades is likely to diminish over time, we expected the degree of distortion to peak soon after episodic retrieval fails, and to remain constant or to diminish after that, depending upon the individual's continuing motivational investment in his or her grades.

We also compare accuracy and distortions for a free recall and a cued recall testing format. To the extent that freely recalled courses involve more easily retrieved episodic memories than the recall of courses that require more contextual cuing, we expect more errors of commission for the cued than the freely recalled courses. However, absent a systematic difference in the degree of emotional investment for these two types of courses, we expect no systematic difference in the degree of distortion of recalled grades.

We also required participants to rate the degree to which they remember enjoying each recalled course and their degree of confidence in the accuracy of the recalled grade. Because enjoyment is likely to be related to emotional investment, we expected it to be associated with more distortion. We expected metacognitive confidence to be positively related to accuracy of recall but not to degree of systematic distortions of content.

Finally, all investigations of false memories must control the possibility that participants deliberately modify their recalled autobiographical content in order to create a more favorable impression or to create some other desired effect. To discourage participants from biasing their responses in the direction of social desirability, we informed them prior to their recall of grades that we planned to compare their reports to the corresponding grades recorded in their academic transcripts, and we secured their written consent to access their transcripts.

Method

Recruitment of Participants

Two hundred seventy-six male and female alumni of Ohio Wesleyan University who had graduated between one and fifty four years earlier participated in the investigation. We recruited 105 of the participants during their attendance at alumni reunions; 26 of these completed both the free and cued recall portion of the investigation, while the remaining 79 completed only free recall. We recruited an additional 171 participants by mail, and they

represented 38% of those who were invited in this manner. Of these participants, 154 completed both free and cued recall, while 17 completed only free recall.

We targeted our recruitment efforts so as to yield a desirable distribution of participants over the 54 years of the retention interval. Table 1 shows the frequency distribution we obtained. Acceptance and completion rates for participants in various retention intervals did not differ significantly ($p > .05$). Our recruitment efforts focused disproportionately on alumni who majored in psychology, and these alumni represent a disproportionate percentage (53%) of our participants. We decided on this recruiting strategy because we expected that a larger percentage of psychology alumni would agree to serve. Also, we believed that the sampling-bias introduced by their over-representation would be smaller and more assessable than the potential sampling bias attributable to having a smaller percentage of targeted recruits agree to serve.

Procedure

Participants signed a consent form granting the investigator access to their college transcript for the purpose of verifying their grades. Participants were assured that the data would be treated confidentially, and as a reward for participating, they were offered a $3 contribution in their name to their class fund. They were informed that after completing and mailing back the responses to this part of the memory test, they would receive a second part of the test by mail.

Free Recall. Participants who attended alumni reunions were given instructions by an investigator who administered the free recall questionnaires. The remaining participants received the same written instructions by mail. These participants completed the free recall test on their own, and they returned the forms by mail.

The free recall questionnaire included a list of all subject-matter areas taught by the college when participants had attended there. Participants were asked to identify all subject-matter areas in which they had taken only one or two courses. For each content-area they identified, they listed the name of the course(s) and the grades they had received in each course. They were asked to omit any plus or minus signs from grades. We decided against including these because the college policy regarding the recording of plusses and minuses had been inconsistent during the fifty-four year period. Participants also indicated on five-point scales their degree of confidence that the recalled grade was correct and their degree of enjoyment of that course.

Next, participants listed, in any order, all of the courses they recalled having taken in their major, and they gave confidence ratings for the recalled grades and enjoyment ratings for each listed course. Participants who

completed more than one major were instructed to choose one and to provide the information only for the major they selected.

Cued Recall: For this part of the investigation each participant received a copy of their academic transcript. The transcript listed all courses the participant had taken, arranged in chronological order by semester. The grades that are normally printed on the transcript had been deleted, and participants were instructed to fill into the appropriate space the grade they received in each course, as well as the ratings of confidence of recall and enjoyment of the course.

Results and Discussion

Table 1 shows the number of alumni tested at various retention intervals for the free and cued recall formats. As mentioned earlier, not all participants returned their transcripts; therefore the cued recall data are available only for a subset of participants, and we defer discussion of these data.

Accuracy of Free Recall

We compared the courses and grades listed on the free recall protocols to the corresponding grades recorded on the transcript, and classified responses as correct, or as errors of commission. Errors of omission were recorded when participants failed to list relevant courses from the transcript. Table 2 shows frequencies of recalled grades by actual grades. The diagonal cells of the table show the number of correctly recalled grades; all other cell entries reflect errors of commission. The table does not show errors of omission; these will be discussed below. Three thousand twenty five of 3967 recalled grades (76%) were recalled correctly; 24% of responses are therefore errors of commission. This finding corresponds fairly closely to the 71% correct recall reported by Bahrick et al. (1996) for high school grades. Greater accuracy of recall in the current data reflects the free recall format which allows errors of omission. The cued recall format of the previous investigation minimized such errors.

Higher grade levels are again recalled with much greater accuracy. Eighty nine percent of As, 72% of Bs and 62% of grades of C are correctly recalled. However, 65% of Ds and 100% of Fs were recalled correctly, and this finding contrasts with much lower percentages reported in the earlier investigation. Because of the small number of Ds and Fs, particularly in the earlier investigation (21 and 1 respectively), the corresponding percentages are unreliable, but they suggest the possibility of a U shaped relation between grade level and accuracy of recall. U-shaped relationships between polarity of affect and accuracy of retention have been reported previously (Rappaport,

1942; Reisberg, Heuer, McLean, & O'Shaughnessy, 1988), with positively valenced and negatively valenced content remembered better than emotionally neutral content.

We used a hierarchical multiple regression analysis to identify predictor variables that have a significant impact on the variance in correct recall. We used regression analyses rather than ANOVAS because we were not able to manipulate the predictor variables in question. The first column of Table 3 shows the results of the regression analysis for recall accuracy. It is apparent that the retention interval, recency of rehearsals, average grades of participants, and subjective ratings of confidence significantly increased the accounted for variance.

Figure 1 shows the mean percent of correct recall as a function of the retention interval and the level of GPA of students, the two variables that have the largest impact on accuracy of recall. GPA level was set by dividing participants into terciles based on the GPA reported on their transcripts. The high GPA group had a mean of 3.64 (*SD* = .19), the medium group had a mean of 3.08 (*SD* = .16), and the low group had a mean of 2.52 (*SD* = .20). Students who earn higher grades recall their grades much more accurately, and accuracy declines with the retention interval. The effect of GPA is in accord with the finding of the earlier investigation. Students with the lowest GPA are again the least accurate. This finding may reflect in part the lower accuracy of generic self assessments of poor performers as reported by Dunning, Johnson, Ehrlinger and Kruger (2003). Declining accuracy of recall during the retention interval was not significant in the earlier investigation, but results of the current investigation are based upon a much longer interval. It is also important to note that accuracy declines during the retention interval because participants recall fewer course grades, while the number of errors of commission remains relatively stable. As a consequence, errors of commission become a progressively larger portion of the total number of recalled grades for longer retention intervals.

Errors of omission were minimal in the earlier investigation of high school grades because the cued recall testing format made such errors very unlikely. The format specified the courses for which grades were to be recalled, e.g., English, mathematics, and foreign languages prompted by year in high school. In contrast, the testing format in the first part of the present investigation provided fewer cues; i.e., the content domains in which the participant had taken courses were not specifically identified. The dearth of specific cues, combined with the greater diversity of college versus high school courses yielded a large number of errors of omission in the current study. Figure 2 shows the increase of omission errors as a function of the retention interval, along with a comparable function for errors of commission and the complementary function, accuracy of recalled responses. The omission

data conform to traditional, negatively accelerated retention functions, and they contrast with functions we have reported in other investigations of highly over-learned memory content. These latter exhibit long periods of stabile retention that we labeled permastore content (Bahrick, 1984c).

The results of a hierarchical regression analysis of the number of errors of omission are shown in column 2 of Table 3. The analysis shows that the retention interval and the level of grades received account for most of the variance. Omission errors increase during the retention interval, and they are more frequent for participants who received low grades. Thus the level of grade is a significant predictor of the accuracy of recalled responses as well as errors of omission.

Distortion of Free Recall

Distortions of recall are reflected by the degree of asymmetry of errors of commission. The previous investigation showed that 82% of errors of commission were in the inflationary direction; i.e., the recalled grade was higher than the actual grade. The asymmetry ratio is obtained by dividing the number of inflationary errors by the total number of errors of commission. Therefore, a ratio of .50 corresponds to symmetry, or the absence of systematic distortion, and ratios below .50 designate a tendency to distort in the deflationary direction. No asymmetry of errors is possible for actual grades of A or F because these grades can only be distorted in one direction. Our calculations of asymmetry are therefore based upon errors of commission for actual grades of B, C, and D. Based on the data in Table 2, 621 of 766 commission errors were inflationary, yielding an overall asymmetry ratio of .81, which is quite close to the .82 ratio reported for the previous investigation.

The distribution of grades recalled for actual grades of B permits the most valid/unbiased assessment of asymmetry in the sense that it is nearly unaffected by erroneous recall of grades of D and F. These latter grades occur with such low frequency that many participants would be able to rule them out by supplementing their specific grade memory with a categorical rather than just a probabilistic generic memory, i.e., the memory that they never received any grades of D or F. The asymmetry ratio based on the distribution of recalled grades for actual grades of B is .75, suggesting that categorical generic memory may contribute somewhat to overall distortion of recall.

We performed a hierarchical regression analysis to identify predictor variables that significantly affect the asymmetry ratio. Column 3 of Table 3 shows the results. Neither the retention interval nor the distinction between students who majored in psychology versus those who majored in other subjects have a significant effect on the degree of asymmetry of errors. The GPA of students and their ratings of subjective enjoyment of the course are the

major predictors. Figure 3 shows that grade distortion is higher for participants who have higher grades. In addition, participants in the highest tercile of enjoyment ratings of their courses had a mean asymmetry ratio of .842 (*SD* = .280); those in the lowest tercile had an asymmetry rating of .724 (*SD* = .335). The analyses confirm the earlier findings that better students recall their grades more accurately (see Figure 1), but their errors show more systematic distortion in the inflationary direction.

Goldsmith, Koriat & Pansky (2005) found that at delayed testing, individuals maintained general accuracy by adjusting the grain size of reported content to compensate for their loss of memory for details of content. This strategy is not available when the grain size of content (e.g., grades in individual courses) is made invariant by the format of the investigation. However, supplementing recall of a specific grade with relevant generic knowledge, e.g., the knowledge that I was a good student, can be regarded as an adjustment analogous to a strategy of increasing grain size when detailed knowledge is no longer accessible.

The effect of recalling grades in major versus non major courses on accuracy and on distortion of recall is a within subject variable that we did not include along with the between subject variables examined in the respective regression analyses. Two dependent t-tests indicated that this variable had no significant effects on either accuracy or distortion of recall ($p > .05$). Figure 4 shows the data for accuracy of recall.

Accuracy of Cued Recall

Table 4 shows the distribution of actual versus recalled grades for courses recalled in the cued recall format but omitted in the free recall format. The data confirm the major findings of the free recall data in Table 2. Overall, only 61% of recalled grades are correct (vs. 76% for the free recall data), but higher grades are again recalled more accurately than lower grades (A=72%, B=63%. C= 45% D= 28%). Errors of commission are again inflationary, with an overall asymmetry ratio of .76. versus .81 for the free recall data. In accord with our expectation, recall is less accurate for content that required additional cuing for retrieval. We expected no systematic differences in the degree of distortion between the free recall and the cued recall format, and the observed lower degree of distortion for cued recall may reflect diminished emotional investment in the less accessible content.

A hierarchical regression analysis of accuracy of the cued recall data is shown in column 1 of Table 5, and the results are consistent with the findings of the free recall analysis. The retention interval, grade level and confidence ratings significantly affect accuracy of recall. In contrast with the free recall testing format, errors of commission increase with the retention interval while the number of errors of omission is negligible.

The regression analysis of the asymmetry ratio is shown in column 2 of Table 5. Consistent with the results for free recall, the ratio is most strongly related to grade level and the enjoyment ratings. Individual differences based upon gender or psychology versus other majors have no significant influence on the asymmetry ratio, and the effect of confidence ratings is very small and only marginally significant.

As stated earlier, students with higher GPA recall their grades with greater accuracy, but their errors show more systematic inflationary distortion. This latter finding conflicts with a dynamic replacement theory of distortion. If systematic distortions occurred because emotionally negative content is actively eradicated and displaced by emotionally positive content, there would be a positive correlation between the number of errors of commission and the degree of distortion. However, this correlation is near zero for free recall (.049, $p > .05$) and negative for cued recall (-.167, $p = .068$). Moreover, groups of individuals with the fewest errors of commission show the greatest systematic distortion. The supplementary theory assumes that losing and distorting memory content are relatively independent processes. For the free recall data, forgetting is primarily reflected by the increasing number of errors of omission, while the number of errors of commission remains relatively constant during the long retention interval (please see Figure 2). It is the asymmetrical distribution of commission errors that defines distortion, and the degree of asymmetry does not change significantly during the retention interval.

The supplementary theory attributes asymmetry to two factors: 1) relevant generic memories used when the specific episodic content is not retrieved, and 2) the tendency to supplement lost episodic content with emotionally gratifying content if that content is compatible with the available generic knowledge. Good students are more likely to commit inflationary errors in recalling their grades because their generic memory content is more compatible with errors in that direction. Poor students also distort content in an inflationary direction, but that tendency is attenuated by their less compatible generic memories.

The above rationale explains the current findings reasonably well, but the findings do not fully correspond to our predictions and leave a number of questions unanswered. In accord with expectations and with our previous findings, higher grades are recalled much more accurately than lower grades, but there is some indication that very low grades may also be more memorable, and our relevant data are inconclusive in that regard. In accord with our expectations, the degree of grade distortion peaks during the first few years of the retention interval and does not change significantly after that. As expected, distortion also increases with the emotional investment in the course grade, as evidenced by the relation between the degree of distortion and ratings of course enjoyment. Counter to our

expectations, however, the degree of distortion does not differ significantly for courses in the participant's major versus other courses.

Although the overall distortion ratio of .81 is highly significant for free recall, 12% of participants distort memory of their grades in a negative direction; i.e.; their recalled grades are lower than their actual grades (14% for cued recall). The data in Table 6 suggest that men are more likely than women to distort in the negative direction, that students with low GPAs are more likely to do so, and that the probability seems to remain relatively stable throughout the 50 year retention interval. The number of individuals who show negative distortion of recall is too small to reach firm conclusions regarding these trends. It also seems likely that uninvestigated personality variables associated with depression or low self esteem may contribute to distortions of recall in the negative direction. .

Concluding Comments

Cognitive memory research has focused on the nature of encoding and retrieval processes for more than 30 years, and issues related to self induced, motivated distortions of memory content have received relatively little attention. Not withstanding this differential emphasis, the past decade has produced important neurological and cognitive advances relevant to distortions of memory and to the effects of emotion on memory. It is well established now that emotion enhances the perceived vividness and in most cases the accuracy of memory (Reisberg & Heuer, 2004) and that the amygdala is critically involved in that enhancement (Adolphs, Cahill, Schul & Babinsky, 1997; Cahill, Babinsky, Markowitsch & McGaugh, 1995; Phelps & Anderson, 1997). It has also been shown that experimental interventions can create false or distorted memories in children as well as adults (e.g.,Bruck & Ceci, 1999; Lindsay, Hagen, Read, Wade & Garry, 2004; Loftus, Coan & Pickrell 1996) and that individuals modify memory of their own past reactions and emotions so as to bring them into better accord with their current feelings and beliefs (Ross, 1989; Levine, 1997).

In contrast, the dynamics of self-induced, systematic distortions of emotionally valenced memory content are relatively unexplored. From the beginning, the focus of memory research has been on quantifying losses of discrete units of content in accord with what Koriat and Goldsmith (1994) have called the storehouse metaphor. Although it was recognized early on that systematic changes of content are pervasive aspects of reconstructing emotionally valenced autobiographical memory, research exploring what Koriat and Goldsmith called the correspondence metaphor has been scant. Methodological problems have been the primary reason for this neglect. Methods appropriate for investigating autobiographical memory reconstructions are likely to offer less control than

traditional methods used to explore quantitative losses of content. This remains true because the emotional conditions driving distortions of autobiographical content are difficult to manipulate or control in the laboratory, and autobiographical content most likely to be distorted is often unique to the individual and experienced in naturalistic settings that are difficult to monitor, verify or replicate.

The present investigation deals with these obstacles by means of regression analyses of a large, naturalistic database. These analyses yield inferences regarding the effects of emotion and motivation on distortions of a fairly comparable memory content. We find it encouraging that the regression analyses account for a sizable portion of observed variance of both the accuracy and the degree and direction of distortions of the recalled content. The findings support a supplementary theory of memory distortion that contrasts with more traditional dynamic replacement theories. The supplementing theory stipulates that losses of content and distortions of content are relatively independent processes, that distortions occur during reconstruction after episodic retrieval fails, that reconstructions are based upon relevant generic memories, and that the distortions are in a direction that generates more gratifying content than the veridical content. However, approximately 12% of participants in the present and the previous investigation distorted memory content in the opposite direction; i.e., the retrieved content appears to be less gratifying than the veridical content. Future research must address the causes of these distortions. The methodological obstacles to the study of self-induced memory distortions are formidable, but not insurmountable. Understanding the dynamics of memory distortions is essential for developing an inclusive theory of memory, and for evaluating forensic testimony based on autobiographical recollections as well as client's reports in psychotherapy.

References

Adolphs, R., Cahill, L, Schul, R. & Babinsky, R. (1997). Impaired declarative memory for emotional stimuli following bilateral amygdala damage in humans *Learning and memory, 4,* 291-300.

Bahrick, H.P. (1984a) Associations and organization in cognitive psychology: A reply to Neisser. *Journal of Experimental Psychology: General, 113,* 36-37.

Bahrick, H.P. (1984b). Replicative, constructive and reconstructive aspects of memory: Implications for human and animal research. *Physiological Psychology, 12* , 53-58.

Bahrick, H.P. (1984c). Semantic memory content in permastore- 50 years of memory for Spanish learned in school. *Journal of Experimental Psychology: General, 113,* 1-29.

Bahrick, H. P. (1996). The relation between reproductive and reconstructive processing of memory content. *Brain and Behavioral Sciences, 19,* 191.

Bahrick, H.P. (1997). Loss and distortion of autobiographical memory content. In C. Thompson, D. Herrmann, D. Bruce, D. Read, D. Payne and M. Toglia, (Eds.), *Eyewitness memory: Theoretical and applied perspectives*, Hillsdale, NJ: Erlbaum.

Bahrick, H.P., Hall, L.K. & Berger, S.A. (1996). *Accuracy and distortion in memory for high school grades. Psychological Science, 7,* 265-271.

Bartlett, F.C. (1932). *Remembering.* Cambridge, UK: Cambridge University Press.

Bruck, M. & Ceci, S.J. (1999). Children's suggestibility. *Annual Review of Psychology, 50,* 419-439.

Cahill, L., Babinsky, R., Markowitsch, H.J. & McGaugh, J.L. (1995). The amygdala and emotional memory. *Nature, 377,* 295-6.

Dunning, D., Johnson, K. Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12,* 83-87.

Edwards, A. (1942). The retention of affective experiences-A criticism and restatement of the problem. *Psychological Review, 49,* 43-53.

Freud, S. (1962). *The standard edition of the complete works of Sigmund Freud* (Vol. 3, J. Sprachey, Trans.). Toronto: Hogarth (Original work published 1899).

Goldsmith, M., Koriat, A. & Pansky, A. (2005). Strategic regulation of grain size in memory reporting over time. *Journal of Memory and Language, 52,* 505-525.

Guy, S.C. & Cahill, L. (1999). The role of overt rehearsal in enhanced conscious memory for emotional events. *Consciousness and Cognition, 8,* 114-122.

Hamman, S.B. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Science, 5,* 394-400.

James, W. (1890). *The principles of psychology* (Vol.1). New York: Holt

Koriat, A.,& Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123,* 297-315.

Levine, L.J. (1997). Reconstructing memory for emotions. *Journal of Experimental Psychology: General, 126,* 165-177.

Lindsay, D.S., Hagen, L., Read, J.D., Wade, K.A., & Garry, M. (2004). True Photographs and False Memories. *Psychological Science, 9,* 6-10.

Loftus, E.F. (1982). Memory and its distortions. In A.G. Kraus (Ed.), *G.Stanley Hall lectures* (pp. 119-154). Washington, DC: American Psychological Association.

Loftus, E.F., Coan, J.A. & Pickrell, J.E. (1996). Manufacturing false memories using bits of reality. In L.M. Reder (Ed.), I*mplicit memory and metacognition (*pp. 195-220). Hillsdale, NJ: Lawrence Erlbaum.

Neisser, U. (1981). John Dean's memory: A case study. *Cognition, 9,* 1-22.

Phelps, E.A. & Anderson, A.K. (1997) Emotional memory: what does the amygdala do? *Current Biology, 7,* R 311-R 314.

Rappaport, D. (1942). *Emotions and memory.* New York: International Universities Press.

Reisberg, D., Heuer, F., McLean, J., & O'Shaughnessy, M. (1988). The quantity, not the quality, of affect predicts memory vividness. *Bulletin of the Psychonomic Society, 26,* 100-103.

Reisberg, D., & Heuer, F. (2004). Memory for emotional events. In D. Reisberg and P. Hertel (Eds.) *Memory and emotion,* (pp. 3-41). New York: Oxford University Press

Roediger, H.L., & McDermott, K.B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21,* 803-814.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review, 96,* 341-357.

Wulf, F. (1922). Über die Veränderung von Vorstellungen (Gedåchtnis und Gestalt). (Concerning changes of imagery (memory and Gestalt)). *Psychologische Forschung, 96,* 341-357.

Author Note

This research was supported by grant number RO1AG19803 from the National Institute of Aging. We thank Mindy Baker an Ann Daunic for editorial comments and Cathy O'Toole and Cameron Hewitt for recording and analyzing data. Correspondence concerning this article should be addressed to H.P. Bahrick, Department of Psychology, Ohio Wesleyan University, Delaware, OH 43015 (e-mail: hpbahric@owu.edu).

Table 1

Number of Participants with Free Recall and Cued Recall Data as a Function of Retention Interval in Years

Retention Interval in Years	Free Recall *n*	Cued Recall *n*
0 to 4	54	40
4.25 to 12.2	56	42
12.25 to 24.5	47	27
24.9 to 39	55	32
39.5 +	64	39
Total	276	180

Table 2

Frequency of Recalled Grades as a Function of Actual Grades on Free Recall

	Actual Grade					
Recalled Grade	A	B	C	D	F	Total
A	1364	282	17	1	--	1664
B	169	1057	281	8	--	1515
C	7	124	510	32	--	673
D	--	1	20	77	--	98
F	--	--	--	--	17	17
Total	1540	1464	828	118	17	3967

Table 3

Summary of Hierarchical Regression Analyses for Free Recall

		Criterion Variable		
Step	Predictor Variables	Percent Correct on Recalled Courses	Percent Omission Errors	Mean Asymmetry
1	Retention interval (Retention interval)^2 (Retention interval)^3	R = .421 F(3, 199) = 14.30***	R = .607 F(3, 200) = 38.91***	R = .144 F(3, 158) = 1.12
2	Months since last viewed transcript	R = .460 F(1,198) = 8.50**	R = .618 F(1, 199) = 4.25*	R = .219 F(1, 157) = 4.48*
3	GPA on all courses GPA on courses in major GPG on nonmajor courses	R = .624 F(3, 195) = 18.90***	R = .669 F(3, 196) = 7.85***	R = .443 F(3, 154) = 9.44***
4	Gender Majored in psychology or other	R = .625 F (2, 193) = .32	R = .671 F(2, 194) = .49	R = .461 F(2, 152) = 1.59
5	Mean confidence rating	R = .642 F (1, 192) = 6.87**	R = .674 F(1, 193) = 1.45	R = .466 F(1, 151) = .98
6	Mean enjoyment rating	R = .643 F(1, 191) = .65	R = .674 F(1, 192) = .001	R = .530 F(1, 150) = 13.20***

Note: Retention interval was expressed as the natural logarithm of the number of months since graduation plus 1.

Each F ratio evaluates the change in R^2 from the preceding step of the analysis. * $p < .05$ ** $p < .01$ *** $p < .001$

Table 4

Frequency of Recalled Grades as a Function of Actual Grades on Cued Recall

	Actual Grade					
Recalled Grade	A	B	C	D	F	Total
A	307	128	21	1	--	457
B	108	388	156	9	1	662
C	9	95	151	20	--	275
D	--	1	8	12	1	22
F	--	--	1	--	4	5
Total	424	612	337	42	6	1421

Table 5

Summary of Hierarchical Regression Analyses for Cued Recall

		Criterion Variable	
		Percent Correct on	
Step	Predictor Variables	Recalled Courses	Mean Asymmetry
1	Retention interval	R = .386	R = .157
	(Retention interval)^3	F(2, 128) = 11.23***	F(2, 100) = 1.26
2	Months since last viewed	R = .399	R = .174
	transcript	F(1,127) = 1.44	F(1, 99) = .57
3	GPA on all courses	R = .661	R = .413
	GPA on courses in major	F(3, 124) = 20.44***	F(3, 96) = 5.42**
	GPG on nonmajor courses		
4	Gender	R = .666	R = .429
	Majored in psychology or other	F (2, 122) = .67	F(2, 94) = .77
5	Mean confidence rating	R = .685	R = .466
		F (1, 121) = 6.03*	F(1, 93) = 3.90 (p = .051)
6	Mean enjoyment rating	R = .686	R = .567
		F(1, 120) = .05	F(1, 92) = 14.12***

Note: Retention interval was expressed as the natural logarithm of the number of months since graduation plus 1. Each F ratio evaluates the change in R^2 from the preceding step of the analysis. * $p < .05$ ** $p < .01$ *** $p < .001$

Table 6

Number of participants who inflated or deflated grades of B, C, or D on free recall as a function of gender, GPA level, and retention interval

	Inflated	Deflated
Gender		
Men	56	13
	(70.9%)	(16.5%)
Women	122	13
	(84.1%)	(9.0%)
GPA		
Low	65	14
	(75.6%)	(16.3%)
Medium	59	12
	(73.8%)	(15.0%)
High	54	0
	(93.1%)	(0.0%)
Retention Interval		
0 – 4 years	23	5
	(79.3%)	(17.2%)
4.25 – 12.2 years	39	3
	(83.0%)	(6.4%)
12.25 – 24.5 years	34	3
	(82.9%)	(7.3%)
24.9 – 39 years	41	8
	(78.8%)	(15.4%)
39.5 years or more	41	7
	(74.5%)	(12.7%)

Note: Twenty participants had the same number of inflationary and deflationary errors. They are included in the calculation of the percentages in each row.

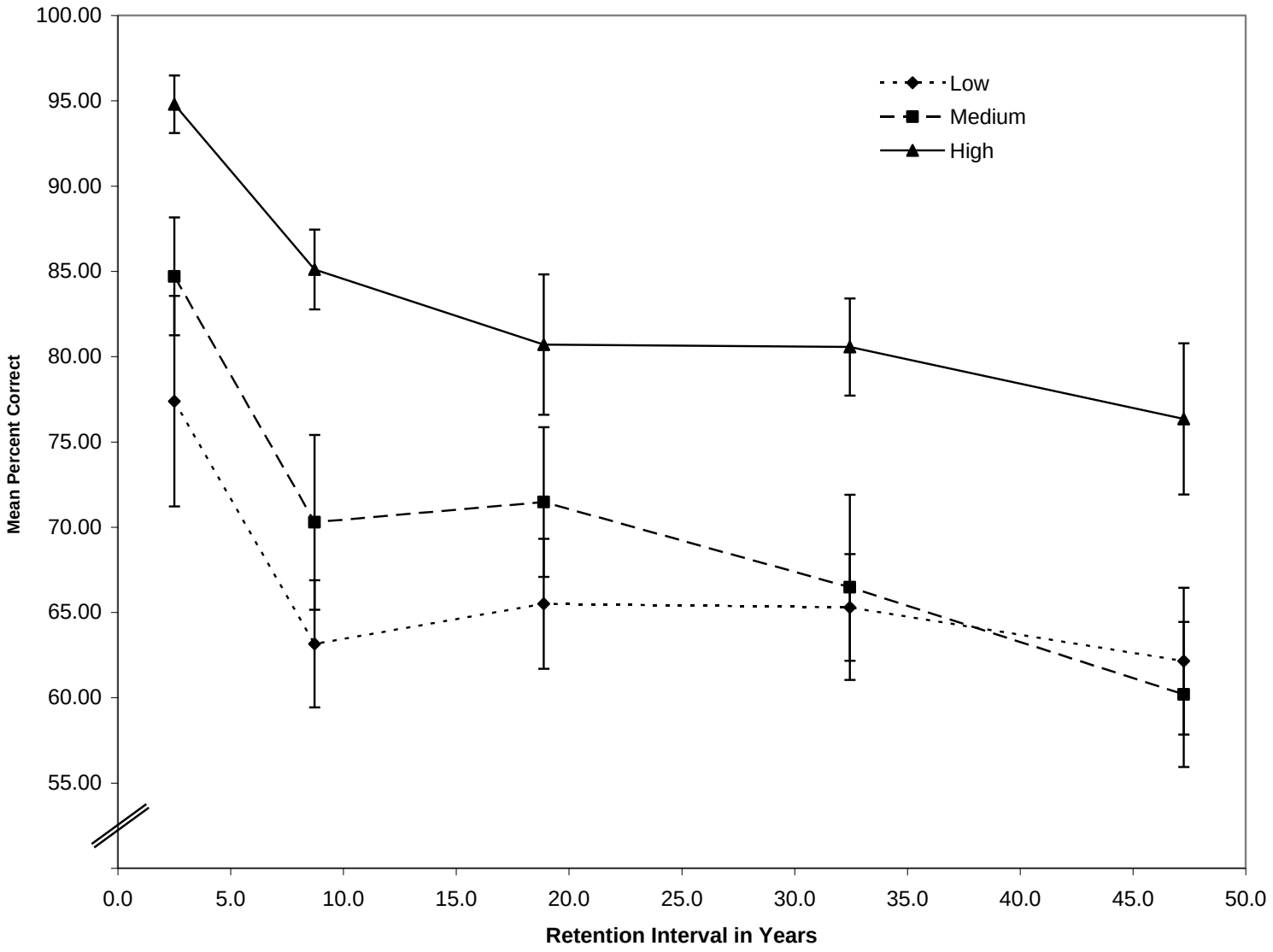
Figure Caption

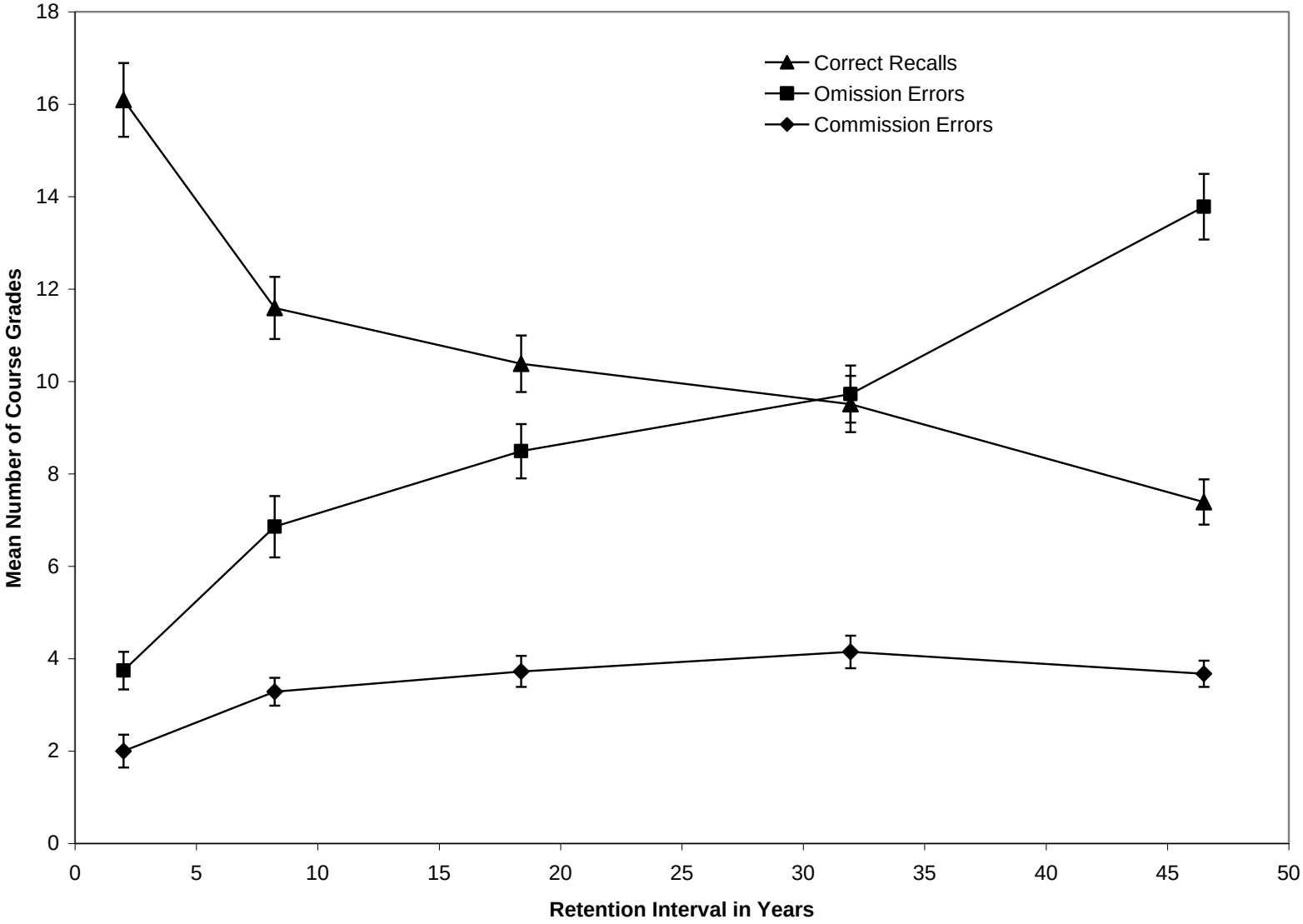
Figure 1. Mean percent correct on free recall as a function of level of GPA and retention interval.

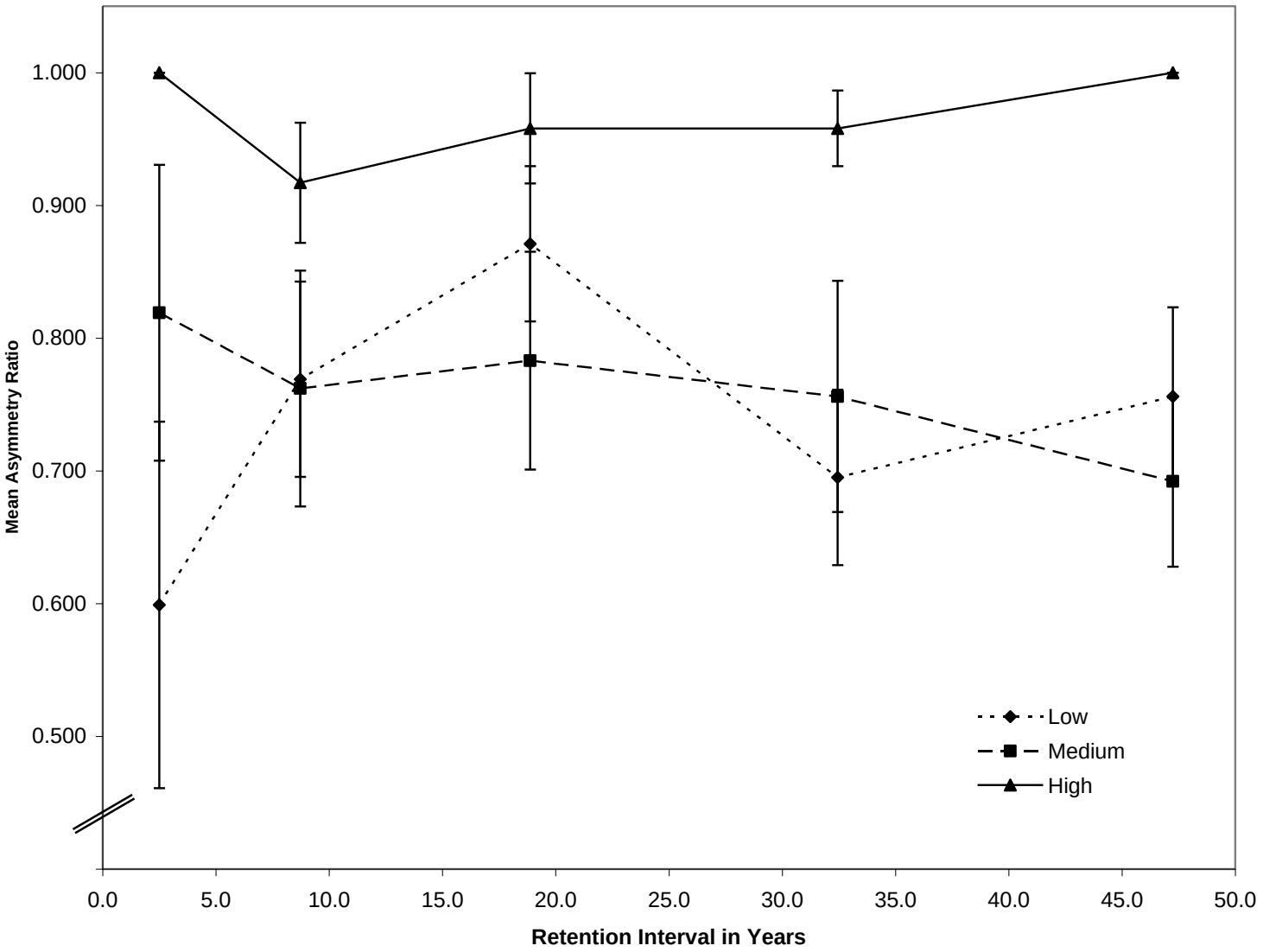
Figure 2. Mean number of correctly recalled grades, omission errors, and commission errors as a function of retention interval on free recall.

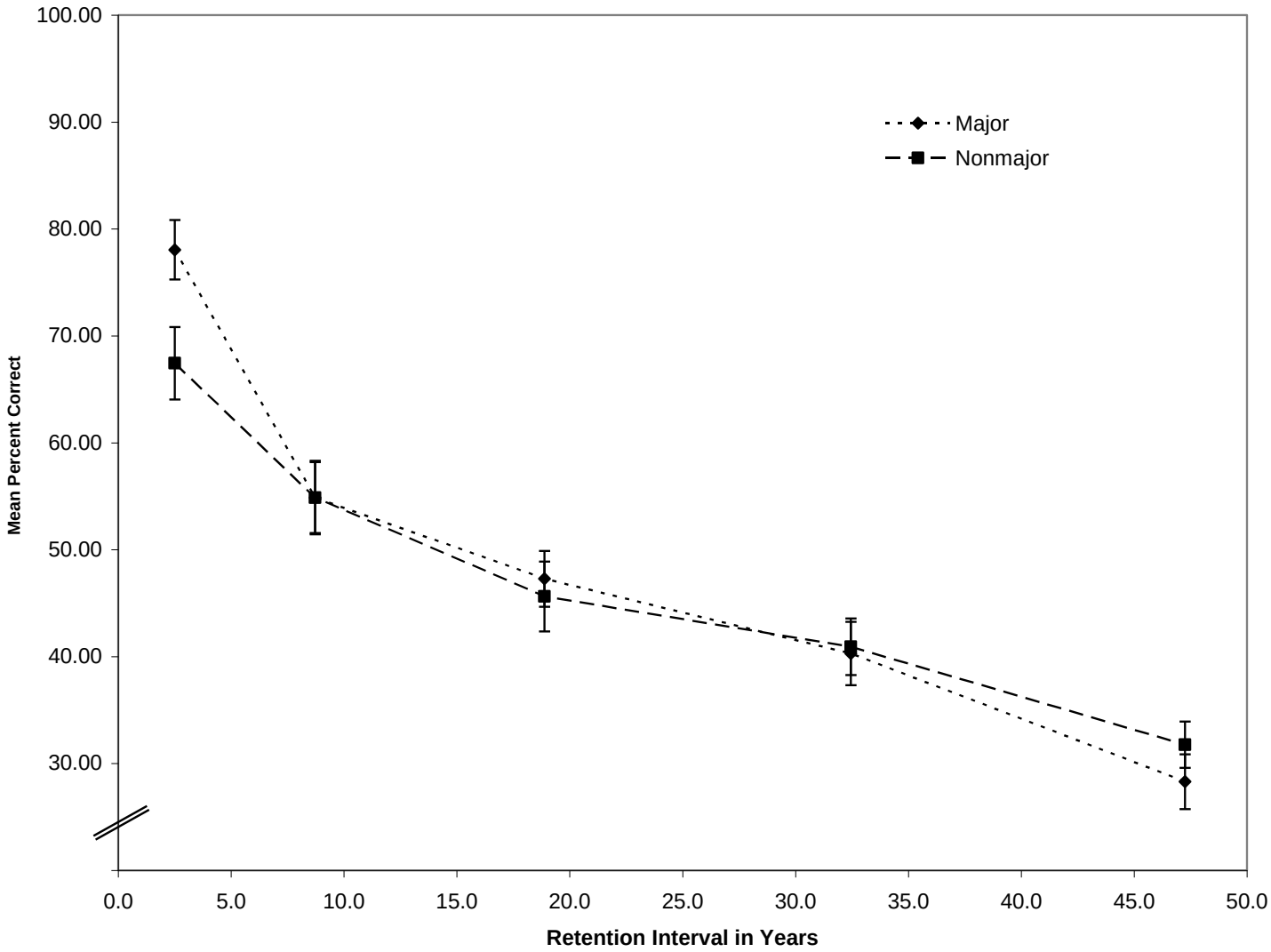
Figure 3. Mean asymmetry ratio on free recall as a function of level of GPA and retention interval.

Figure 4. Mean percent correct on free recall for major versus nonmajor courses.


100.00
95.00
90.00
85.00
80.00
75.00
70.00
65.00
60.00
55.00
Mean Percent Correct
Low
Medium
High
0.0
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
45.0
50.0
Retention Interval in Years


Correct Recalls
Omission Errors
Commission Errors
Mean Number of Course Grades
Retention Interval in Years
0
2
4
6
8
10
12
14
16
18
0
5
10
15
20
25
30
35
40
45
50


1.000
0.900
0.800
0.700
0.600
0.500
Mean Asymmetry Ratio
0.0
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
45.0
50.0
Retention Interval in Years
Low
Medium
High


100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
Mean Percent Correct
Major
Nonmajor
0.0
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
45.0
50.0
Retention Interval in Years