LEVELS OF PROCESSING

Craik, F. I. M. & Tulving, E. (1975)
In groups answer following questions...

- Using your knowledge of Craik & Tulving (1975)...
- List any vocabulary that is not understood
- What was the primary research question?
- What was the hypothesis?
- What is the design (between vs within)?
- For experiment 2, what was the IV (what are the levels)?
- For experiment 2, what were the DVs?
- For experiment 2, what types of statistics were completed?

- In the experiment you completed for lab 2...
- What is the IV (what are the levels)?
- What are the DVs?
Introduction

- **Research questions**
  - Examine “the processes involved in learning and remembering”
  - “How should we conceptualize learning and retrieval operations...?”
  - Why is info remembered or not?
  - What is required for good memory performance?

- **Prior knowledge**
  - Hyde (1973); Schulman (1971): when Ss analyze meaning of words in list subsequent recall is excellent
  - Result found for intentional or incidental procedures
  - Bobrow & Bower (1969): good retention for sentences
  - “Orienting task requiring semantic or affective judgments led to better memory performance than tasks involving structural or syntactic judgments.”
  - Why? What are the implications?
Hypotheses and general method

- The “durability of the memory trace is a function of depth of processing.”
- “Stimuli that are attended to, fully analyzed, and enriched by associations or images yield a deeper encoding of the event, and a long-lasting trace.”
- “Memory performance would vary systematically with the depth of processing.”
- 10 experiments on LOP (“levels of processing”) on college students

General method:

- Study phase
  - “orienting task” or “initial test” or “perceptual-reaction time test”
  - Ss asked question to induce word processing at certain level
  - Word presented for 200ms; recorded y/n decision

- Test phase: memory test
  - Unexpected (“incidental”) test
  - Free recall, cued-recall, or recognition test
Experimental design

- What type of design is used?
  - Within-Ss design

- What are the IVs?
  - Study phase: Level of processing (3+ levels)
  - Study phase: Response type (correct answer) (2 levels)

- What are the DVs?
  - Study phase: accuracy
  - Study phase: response time (latency) (only accurate trials) (calculated medians for each Ss)
  - Test phase: accuracy

- Hypothesis
  - “Memory performance would vary systematically with depth of processing”
Experiment 1: Method

- **Study phase:**
  - 5 types of questions (levels of encoding):
    - 1. Is there a word present?
    - 2. Is the word in capital letters?
    - 3. Does the word rhyme with _____?
    - 4. Is the word in the category _____?
    - 5. Would the word fit in the sentence _____?
  - **Details:**
    - Common 2-syllable nouns 5-7 letters
    - 40 trials: 4 words for each of 10 conditions (5 levels x y/n)
    - Counterbalanced words (between-Ss)
  - **Method:**
    - Question read to Ss; after 2s word shown; Ss respond y/n (see word 200ms)

- **Test phase:**
  - Recognition test
  - 80 trials: 40 original words & 40 similar new (distractor) words
  - No time limit for each response
Experiment 1: Results

- RT increases w/ depth of processing
- RT = for y/n answers
- Acc increased w/ depth of processing
- Acc better for yes responses from study than no responses (Note: ceiling effect for sentence level)

Conclusion:
- Deeper processing leads to better retention
- Why? Just b/c longer study time? Why are ‘yes’ responses at study higher acc?

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>INITIAL DECISION LATENCY AND RECOGNITION PERFORMANCE FOR WORDS AS A FUNCTION OF INITIAL TASK (EXPERIMENT 1)</td>
</tr>
<tr>
<td>Response type</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Response latency (msec)</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Proportion recognized</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
</tbody>
</table>

Note: N = 20
Experiment 2: Method

- **Replication:**
  - Simplify design and remove ceiling effect

- **Study phase:**
  - 3 levels of encoding of 5-letter common nouns
    - Physical: Is word in uppercase or lowercase?
    - Phonemic: Does word rhyme with ____?
    - Semantic: Does word fit into sentence ____?
  - Details
    - 60 questions (10 y/10n for each level)
    - Words counterbalanced (between-Ss)

- **Test phase:**
  - Recognition test: 180 words
  - 60 original words (old); 120 distractors (new)
Experiment 2: Results

Figure 1. Initial decision latency and recognition performance for words as a function of the initial task (Experiment 2).

**Note:** N = 24

- Replicated Exp 1 results
- Is effect due to just processing time or is level of analysis the critical factor?
- Hypothesis: effect is due to depth of processing

![Latency vs Level of Processing](image1)

![Proportion Recognized vs Initial Decision Time](image2)

**Figure 2.** Proportion of words recognized as a function of initial decision time (Experiment 2).

- Longer RT for deeper processing
- Memory increases for deeper processing
- Y response at study increase memory
Other experiments: Replication with recall tests and intentional instructions

- **Exp 3: Unexpected recall test (N = 20)**
  - ½ words presented 2x to boost performance (same type of encoding question) (reduce floor effect)
- **Exp 4: Expected recall test (N = 20)**
Other experiments: Use shallow task that takes longer than deep task

- Exp 5: \( (N = 24) \)
  - Non-semantic: Does word fit pattern of vowels and consonants (Is “brain” characterized as CCVVC?)
  - Semantic: Does word fit into sentence?
  - Surprise recognition test
  - Result: levels of processing (not time) determines memory performance

<table>
<thead>
<tr>
<th>Response type</th>
<th>Level of processing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Structure</td>
</tr>
<tr>
<td>Response latency (sec)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.70</td>
</tr>
<tr>
<td>No</td>
<td>1.74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Proportion recognized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
</tbody>
</table>
Other experiments: Questions provide same level of depth for y/n responses

- **Exp 6: (N = 16)**
  - Is word greater than or less than object (size, length, width, height, weight, temp, sharpness and value)
  - Ex. Is ___ bigger than a chair?
  - Ex. Is ___more valuable than $10?
  - Regardless of y or n answer, same amount of elaboration required
  - Recall test

Recall Results for Exp 6
- Yes responses: .36
- No responses: .39
- Non-significant diff

Conclusion: not response, but richness of elaboration determines memory performance
Exp 7: (N = 20)
- Sentence task: simple, medium, or complex
  - “She cooked the ______” (simple)
  - “The _____ is torn.” (medium)
  - “The _____ frightened the kids.” (hard)
- 60 words (20 for each level)
- Surprise free recall
- Surprise cued-recall (given sentence)

Results:
- No effect of sentence type on no responses
- Complexity of semantic encoding large effect on yes responses
Other experiments: Examine if effect is due to having unique retrieval cues

- Exp 8: (N = 32; by grp = 11, 11, 10)
  - Is better memory for deep condition b/c there are unique retrieval cues for each word (opposite for shallow condition)?
  - 3 grp of Ss: vary # of trials of each question type
    - If use fewer trials (4 vs 40) does recognition improve for shallow cond’s
  - Result: No/Little effect of # of different types of questions

<table>
<thead>
<tr>
<th>Experimental condition</th>
<th>Case</th>
<th>Rhyme</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Group 1</td>
<td>2</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Group 2</td>
<td>8</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>Group 3</td>
<td>20</td>
<td>20</td>
<td>2</td>
</tr>
</tbody>
</table>

Design: Number of trials per condition

Proportion recognized

- Group 1
  - Set size 4: .50 .36
  - Set size 16: .51 .40
  - Set size 40: .49 .43

- Group 2
  - Set size 4: .73 .47
  - Set size 16: .66 .54
  - Set size 40: .90 .70

- Group 3
  - Set size 4: .88 .70
  - Set size 16: .95 .64
  - Set size 40: .91 .68
Other experiments: Examine effect if loosen experimental control

- Goal: loosen experimental rigor (control)
- Classroom demonstration
- Exp 9 (N = 12)
  - Informed of the purpose of the study; used intentional recognition test
  - Ss tested in 1 group; words shown for 6s each; Ss recorded own answers; RT not measured
  - Results similar to Exp 2 despite loose experimental rigor
  - Ss had 6s to study words but still didn’t recognize words from shallow condition well (42%)

<table>
<thead>
<tr>
<th>Response type</th>
<th>Case</th>
<th>Rhyme</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>.23</td>
<td>.59</td>
<td>.81</td>
</tr>
<tr>
<td>No</td>
<td>.28</td>
<td>.33</td>
<td>.62</td>
</tr>
<tr>
<td>2nd study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>.42</td>
<td>.65</td>
<td>.90</td>
</tr>
<tr>
<td>No</td>
<td>.37</td>
<td>.50</td>
<td>.65</td>
</tr>
</tbody>
</table>
Other experiments: $ to attempt to increase performance on shallow condition

- **Exp 10 (N = 36; 12 per grp)**
  - Correct recognition for 1c, 3c, or 6cents for each of the levels
    - Grp 1: 1c, 6c, 3c for case, rhyme, category, respectively
    - Grp 2: 3c, 1c, 6c respectively
    - Grp 3: 6c, 3c, 1c respectively
  - **Results:** different rewards had no effect on memory performance!
  - Even though motivated to learn (case level): still worse than rhyme and semantic level

<table>
<thead>
<tr>
<th>Encoding operation</th>
<th>Reward value</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 cent</td>
<td>3 cents</td>
</tr>
<tr>
<td>Case</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>.50</td>
<td>.51</td>
</tr>
<tr>
<td>No</td>
<td>.51</td>
<td>.50</td>
</tr>
<tr>
<td>Rhyme</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>.73</td>
<td>.73</td>
</tr>
<tr>
<td>No</td>
<td>.53</td>
<td>.50</td>
</tr>
<tr>
<td>Category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>.93</td>
<td>.89</td>
</tr>
<tr>
<td>No</td>
<td>.72</td>
<td>.75</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>.72</td>
<td>.71</td>
</tr>
<tr>
<td>No</td>
<td>.59</td>
<td>.58</td>
</tr>
</tbody>
</table>
General discussion

- Experiments 1-4 show “subsequent memory performance is strongly determined by nature of” judgment from study phase
- Experiments 5-10 isolate possible reasons for effect
- “It is abundantly clear that what determines the level of recall or recognition of a word event is not intention to learn, the amount of effort involved, the difficulty of the orienting task, the amount of time spent making judgments about the items, or even the amount of rehearsal the items receive; rather it is the qualitative nature of the task, the kind of operations carried out on the items, that determines retention.”
- What is the theory?
  - More than just “meaningful things are well remembered”
General discussion

Depth of processing
- Stimulus is processed through fixed series of analysis from structure to semantic level
- Level reached determines strength of memory trace
- Problems:
  - Is it really a continuum from structure to semantics?
  - Exp7 finds higher recall for more complex sentences. (How does theory also include levels within levels? )
  - Same effects found for intentional and unintentional learning conditions

Degree of encoding elaboration
- Elaboration based on context given by encoding (study phase) question
- Not processing sequence; depends on elaboration
- Higher retention of yes decision due to integration with encoding context; retrieve entire unit at test

Implications
- Creates new research problems; think about memory in new way
- New question: “what are the encoding operations underlying ‘normal’ learning and remembering?”
Lab 2: LOP experiment

- What are IVs? What are DVs? What type of design?

- Study phase
  - Physical condition: “Does the word have 4 letters?”
  - Rhyme condition: “Does the word rhyme with ____?”
  - Meaning condition: “Does the word fit into this sentence?”
  - 60 words; 20 per condition
  - Responses: 10 ‘yes’ & 10 ‘no’ for each condition
  - 2s to see condition instructions; RT for response

- Test phase
  - Surprise recognition test: y/n response (untimed)
  - 180 words: 60 old studied words; 120 new foils
Lab 2: LOP experiment

- **Type of design?**
  - Within-Ss: every completes all conditions, types of responses, types of test questions

- **What are IVs?**
  - Study condition (question type) — 3 levels (physical, rhyme, meaning)
  - Study responses — 2 levels (y/n)

- **What are DVs?**
  - Study phase accuracy (y/n decision)
  - Study phase RT (only correct trials)
  - Test phase accuracy (old/new decision)
Lab 2: Experimental design

- What are the potential confounds (‘threats to internal validity’)?
  - Time of measurement
  - Fatigue effect
  - Measurement error
  - Diffusion of treatment (knowledge)
  - Participant effects (bias)
  - Ceiling/floor effects
  - What is “control” grp?
  - What is the counterbalance procedure?

- What are potential external validity concerns?
  - Does sample generalize to population of interest to draw conclusions about (adults)?
  - Does study tell us about real-world memory?
Lab 2: Analyses

- Make graphs of what you expect the results of Lab 2 to look like!
- What numbers (means/sd’s) do you need to know?
- What are your hypotheses?
- Begin to think about... if you did a LOP study with a new twist... what would you do??
  - Need ALL the specifics of methodology (# of trials, how you counterbalance, etc.)
  - Choose a partner for presentation
  - Get ideas from other LOP or DOP studies - PsycINFO
Lab 2: Timeline

- HW for 11/10: write intro and methods
- Wed, 11/10 in Olin: data analysis
- HW for 11/17: write results
- Wed, 11/17 in Olin: analyses, search for articles
  - HW for 12/1: write discussion (revise all)
  - HW for 12/8: work on presentation
- Wed 12/1: peer review
  - HW for 12/8: revise paper; work on presentation
- Wed 12/8: presentations