Survey of Attention Seminear
John Palmer
Winter 1999

This is a summary of a seminar that covered Hal Pasler's book on attention. This document contains the sylabus, weekly question sets, and a summary of the discussion for each week. Hal participated in the final week's disscussion.

This collection was assembled long after the fact.
23 October 2005
John Palmer

Survey of Attention
Course Syllabus
John Palmer
7 January 1999

Synopsis

We will read and discuss Pashler's recent book on attention. It is a survey of attention in vision, audition, memory and motor control. It is particularly interesting in revealing the common phenomena across this wide range of topics. In my opinion, it is the best book on attention since Broadbent's 1958 landmark text.

General Information

Psychology Graduate Seminar 560P, 2 Credits, CR/NC, SLN 8511
Weekly meetings 4:00-5:30+ PM on Thursdays in 211 Guthrie.
Entry Codes and other information available from John Palmer

E-mail:	jpalmer@u.washington.edu
Phone:	543-0706
Address:	Psychology, Box 351525, U of W

Readings

Pashler, H. E. (1998) *The Psychology of Attention*. MIT Press.
Available at the University Bookstore

Course Structure and Responsibilities

John Palmer will jointly lead the discussion with one or two discussants selected for the week. To help guide discussion, study and discussion questions will be distributed a week ahead. The study questions have answers that can be found in the reading while the discussion questions are open ended. Seminar participants will be expected to read the material in detail before the class meeting. Be sure to understand the study questions and think about the others. Everyone is expected to be a discussant once or twice. Discussants are expected to meet with the Instructor a few days before class and to draft a summary of the class discussion. This summary will be distributed to all participants by e-mail before the next meeting.

Survey of Attention
Schedule of Topics
John Palmer
7 January 1999

Date	**Topic and Readings**
1/7	Organizational Meeting
1/14	Chapter 1: Introduction
1/21	Chapter 2: Selective Attention
1/28	Chapter 3: Divided Attention
2/4	Chapter 4: Attentional Set
2/11	Chapter 5: Capacity and Selection: Theorizing about Attention
2/18	Chapter 6: Central Processing Limitations in Sensorimotor Tasks
2/25	Chapter 7: Attention and Memory
3/4	Chapter 8: Automaticity, Effort, and Control
3/11	Conclusion and Synthesis

Survey of Attention
Questions for the Introduction
John Palmer
7 January 1999

Study Questions

1. How does Pashler characterize the folk psychology of attention? What pitfalls in the use of the term "attention" does he point out?

2. What are the differences between early and late selection theory? What are selection and capacity? (In the chapter, these terms are defined by example only.)

3. How are early and/or late selection theory modified by the suggested alternatives of the controlled parallel theory, attenuation, and capacity sharing?

4. How do filter and monitoring tasks differ? What are examples of each?

Discussion Questions

1. In this chapter, selection and capacity are presented as the primary components of attention. Are there any other components of equal importance that we should consider? If not, does this constrain how one should think about attention?

2. What are the conflicts between the various folk and scientific uses of the term "attention"? How should one use (or not use) the term?

3. Might the controlled parallel theory be a better starting point for the analysis of attention than either early or late selection theory because it reflects the optimal processing strategy? Are there precedents to this theory? Is there a better name?

4. Pashler briefly contrasts the approaches of introspection, information processing, and neuroscience. What are the advantages and disadvantages of each? How can these alternative approaches be compared?

5. While not done in this introduction, would it be useful to define the phenomena relevant to the study of attention? In particular, how can one exclude clearly nonattentional phenomena such as light adaptation or perceptual learning?

6. Pashler comments on the scientific method by emphasizing critical experiments that test among competing theories. How can one clarify his distinction between this kind of theory testing and what he refers to as data fitting?

Survey of Attention
Synopsis of Discussion
Week 1: Introduction
Andrea Hartzler and John Palmer
19 January 1999

Readings

Pashler, H. E. (1998). The Psychology of Attention. MIT Press. Chapter 1.

Palmer, J. (1992). What Defines Attentional Phenomena? Transcript of talk presented at the meeting of the Society of Mathematical Psychology, Stanford, CA, August, 1992.

What are filtering and monitoring tasks?

A filtering task is defined as a task in which a subject attends to a subset of stimuli and ignores other irrelevant stimuli. Filtering tasks are associated with selective attention and are often called selective attention tasks. Examples of filtering tasks include shadowing one channel during dichotic listening and the partial report of cued letters from a visual array. A more colorful example introduced in discussion is to ignore one's mother-in-law while driving.

A monitoring task is defined as a task in which a subject must attend to multiple stimuli. Monitoring tasks are associated with divided attention and are often called divided attention tasks. Examples of monitoring tasks include visual search, simultaneously watching and listening to a television, and simultaneously watching two superimposed videos. A slightly more colorful example introduced in discussion is an instructor listening for one of many students to speak up.

What is selection and capacity?

Selection is the differential processing of different sources of information. Examples of selection discussed include moving the eyes toward one stimulus and away from another, processing sounds from one ear and not the other, and trying to remember the animal words from a list of words.

Capacity is a measure of the amount of information being processed. It can be defined in several different ways. Pashler begins by introducing limited capacity as a measure of "our limited ability to carry out various mental operations at the same time" (p. 2). He has a more complete discussion at the beginning of the Divided Attention chapter. Several comments were made. One was to focus on the definition of unlimited capacity as each component processes being independent of any other concurrent processes. Limited capacity then becomes the degree of dependence. This leaves open how to measure the degree of dependence. Another comment was to raise the idea of supercapacity. This is the possibility that the more mental resources needed, the more become available. Finally, does capacity refer to an entire behavior, a particular stage of processing such as perception, or a single component of a larger processing system. All three usages are common in the literature and are easily confused.

Further questions were raised about how to define capacity more exactly. For example, if one had unlimited capacity, why would it matter whether selection was early or late? Are there different types of capacity that are specific to a task or available under some circumstances? Issues related to selection also resurfaced at this point. Does one necessarily relate unattended information to things one has previously learned? Why do we recognize our name when it is presented in an unattended channel? None of these topics were pursued.

Are there other equally important concepts?

Pashler focused on selection and capacity. But are they alone the fundamental concepts of attention or are other concepts equally important? For example, effort and arousal are not even mentioned. Could these concepts be some of the building blocks for a theory of attention? Alternatively, can effort and arousal be described using capacity and selection? We briefly discussed Kahneman's book which closely related arousal and capacity. Barry Kantowitz critiqued the idea of such a simple connection to aggregate physiological measures. Other suggested candidates of fundamental concepts involved learning and salience.

A different concern was whether selection and capacity can be defined without some explicit statement about the architecture of the information processing system. In the chapter, Pashler implicitly refers to a simple communication channel metaphor following Broadbent. But not everyone was

satisfied with this informal development. Some thought issues of the system architecture were as fundamental as capacity or selection. For example, what are the component processes? How are they related to the stimuli and other experimental measures? Do they create distinct or continuous changes in the representations? Some were unclear about how capacity related to serial versus parallel processing. (Pashler makes his usage clear at the beginning of the divided attention chapter.) While all of these issues can't be raised in an introduction, some introduction to these topics may be on par with selection and capacity.

What are the simple theories of attention from which to begin?

In early selection theory, the selection is assumed to be early in processing. Information from a single stimulus is chosen to be fully processed. Thus, to be fully processed, multiple stimuli must be processed serially. Early selection theory also assumes limited capacity in the more complex aspects of perceptual processing.

In late selection theory, selection is assumed to be late in processing. The idea is that all stimuli are processed in parallel and only at the point of establishing a durable memory is information from a single stimulus selected for further processing into memory. Late selection theory assumes unlimited capacity in perceptual processing. Experiments in which a subject can recognize information presented to an unattended channel are consistent with late selection theory. The Stroop effect is another example of a phenomenon consistent with this theory.

The controlled parallel theory has elements of both early and late selection theories. It assumes that selectivity can happen early if desired, but need not and there are no capacity limits on perception. This combination combines elements of both theories to yield optimal performance: selection to avoid interference and unlimited capacity to allow processing of multiple inputs.

But then, one asks, how does such a controlled parallel model explain phenomena such as the Stroop effect where selection fails? One must assume specific limits on the degree of selection. For example, one may be able to select a region in space but not ignore a word within that region. In addition, one may be able to process both the ink color and the word, but not select between the two resulting color tokens. The resolution of such issues is in the details.

It was suggested that the guided search model of Wolfe is an example of a controlled parallel theory. Indeed, this and other two-stage search theories also combine elements of early and late selection. However, they are not very specific in whether selection is possible for the early parallel processing. This bears further thought.

When should one use the term "attention"?

Care must be taken when using the term "attention". The common sense meaning of attention carries many connotations of which no consensus exists. For example, it is dangerous to assume too early that one is measuring attention in a given experiment. For example, sometimes a set-size effect is taken to be an attention effect by definition. But sometimes set-size effects might be due to other processes other than attention such as eccentricity, crowding or masking.

Another example is the statement, "Is attention necessary for memory?". Such statements are meaningful to the degree that there is a consensus about the mechanisms of attention. Although there may be a consensus of the common sense meaning of attention, there is no such consensus on either the phenomena or mechanisms of attention.

A variety of questions were raised in class regarding the definition of attention. Might attention be a global property of a system rather than a specific property of a component within a system? Would we be better off ignoring the term of attention altogether and instead use narrower concepts such as selection and capacity? How do these issues relate to distinctions between sensory, perceptual, memory, and motor processing? Do the particulars of what one is studying shape your definition of attention?

How can one define attentional phenomena?

This discussion begins from the counterpoint article (Palmer, 1992). This article focused on the phenomena of attention, setting aside issues of mechanism and theory. The article discusses phenomena as sufficient, as necessary or as necessary and sufficient to be considered purely attentional phenomena.

Sufficient properties are the least controversial. Narrowly and

appropriately defined tasks such as cueing in visual search or allocation instructions in dual tasks are clearly examples of attentional phenomena. The use of identical stimuli and tasks leaves only effects initiated by voluntary control. Such narrow definitions are useful to identify laboratory idealizations, but do not help organize the larger world of possible tasks.

At the other end of the spectrum is the effort devise a necessary property. Perhaps attention always involves a control process. This does not mean that attention is equated with control. Thermostats and light adaptation both are examples of control processes and are not attentional. It is OK for a necessary property to also be true of non-attentional processes.

Necessary and sufficient conditions are by far the most difficult. The article suggested a definition based on voluntary control. The obvious challenge to this approach are examples of orienting that appear to be involuntary. In the article, there is an augment that orienting can be modulated voluntarily under some circumstances, but this leads to a painfully phrased definition: "A phenomena is purely attentional if and only if it can be completely regulated by voluntary intentions."

It was questioned whether necessary and/or sufficient conditions were unrealistic? Might it be enough to agree on a number of laboratory idealizations? Palmer argued that an incomplete definition is better than no definition.

What is Pashler's distinction between data fitting and theory testing?

Pashler questions the usefulness of fitting data to complex models. According to Pashler, data fitting can help improve a theory by adding new descriptive data. But, Pashler warns the reader that data fitting does not substitute for testing alternative hypotheses. In order to test a theory, one must identify the properties that are required by a particular model. Without this, one cannot reject the model. The group wanted more examples of data fitting to evaluate these ideas.

On a related topic, Pashler's book minimizes the presentation of formal theory. Palmer quoted him as wanting to avoid premature quantification. But when is quantification premature? Eric Horvitz introduced a related distinction in computer science: researchers are considered either "scruffy" or "neat". Scruffies use what ever they need to to solve a problem; Neats like to define

their problems carefully and prove theorems. Is Pashler a scruffy or a neat? Barry Kantowitz raised a related issue of exposition. Quantitative arguments limit your audience to those that have the tools to understand the argument. Thus, avoiding extensive quantification potentially allows one to reach a larger audience. Does Pashler successfully reach a larger audience? These questions may be worth returning to at the end of the seminar.

Survey of Attention
Questions for the chapter on selective attention
John Palmer
14 January 1999

Study Questions

1. What attributes allow easy selection (e.g. location)? What ones don't (e.g. letter identity)? What criteria are used to make this distinction?

2. In filter tasks, there is little evidence for the processing of the unattended stimulus using direct measures such as recall, but somewhat more with indirect measures such as the Stroop effect, flanker effects, and negative priming. What alternative interpretations of this pattern of results does Pashler describe?

3. How long does it take to shift from one selection criterion to another?

4. What are Pashler's four conclusions? Can you put them in your own words?

5. Why might location may be different than other attributes for visual selection?

Discussion Questions

1. For each of Pashler's last two conclusions, what single study is the best departure point for further analysis? What further work would be desirable?

2. In the discussion of bimodal selection, the issue of multiple simultaneous selection criteria is broached. What is known about this within a modality?

3. There may be a relation between eye movements and visual attention. Is the evidence for a mandatory shift convincing? What other action and sensory systems might be similarly related?

4. The ease of selection must depend on the discriminability of the relevant attributes. For example, selecting one of two easily discriminated locations will be easier than selecting one of two poorly discriminated colors. To better compare selection of different attributes, might one measure the difference in attribute values necessary to obtain a criterion degree of selection? What criterion could be used? What problems does this approach face?

5. In the review of the fate of unattended attributes, Pashler argues that the results are inconsistent with both perfectly selective or completely unselective processing. How might the degree of selectivity be defined and studied?

Survey of Attention
Synopsis of Discussion
Week 2: Selective Attention
Maria Pereverzeva and John Palmer
27 January 1999

Readings

Pashler, H.E. (1998). The Psychology of Attention. MIT Press. Chapter 2.

Smallman, H. S. & Boynton, R. M. (1990). Segregation of basic colors in an information display. Journal of the Optical Society of America, 7, 1985-1994.

What determines the ease of selection?

Under what conditions is selection easy or hard? Filtering is particularly successful when based on attributes of location, color, size and brightness for partial report tasks; and, location and frequency for dichotic listening tasks. Less success was reported for attributes of orientation, letter vs. digit, and vowel vs. consonant.

How does one measure this effect? In Pashler's examples, the ease of selection was measured in terms of the number of shadowing errors, or the advantage of partial report over full report. Another approach was to use search time and was described in the counterpoint article (Smallman & Boynton, 1990).

Lynne Werner suggested that the analysis of selection didn't take advantage of theories of basic sensory processing. For example, the removal of onset and offset cues was described as making selection more difficult (Treisman & Riley, 1969). Such timing changes are expected to make a sound more difficult to localize. One could use experiments on the localization of sound to predict the ease of filtering by location. While Pashler would agree with this general idea, Lynne argued he didn't exploit it for all that it was worth. Others suggested that this kind of argument may be extended to issues of perceptual organization if it is objects and events that are selected rather than more elementary sensory representations. The discussion also touched on comparisons between the role of frequency in audition and location in vision. More generally, this entire discussion emphasizes the need for more explicit

assumptions about the representation upon which selection acts.

Is the ease of selection determined by discriminability?

The discussion turned to the general role of discriminability in determining the ease of selection. Someone asked about the degree to which Smallman and Boynton equating the discriminability of their stimuli. In fact, they picked their basic colors on other kinds of data rather than just discriminability. But, when they compared basic and non-basic colors, they carefully matched the discriminability of the two sets. Thus, while they didn't address discriminability directly, they were concerned about it mediating selection. No one present thought that any of the other studies discussed showed such concern with controlling discriminability. This raises the question of how does one compare selection based on color to selection based on orientation? Perhaps the observed differences in selection are due to the color patches being much more discriminable than the orientations. Indeed, the most discriminable attributes in vision are almost certainly location and color. Perhaps this is the reason that selection is so easy with only those two attributes. Perhaps selection by digit versus letter would be similar to location and color if the level of discriminability was held constant. Of course to make such a comparison, one would have to make the locations or colors very similar.

What is the fate of the unattended information?

The discussion next turned to the fate of unattended information. Two examples were discussed in some detail. The first was Treisman's (1960) shadowing experiment where a prose passage being shadowed switched from the attended ear to the unattended ear. The result is that subjects sometimes followed the prose for at least a brief time rather than shadowing the instructed ear. The early selection account of this result depends on noticing the prose no longer makes sense in the to-be-attended ear and then switching to the other ear while sensory memory still maintains that information. This early selection account seemed plausible but should lead to further testable predictions.

The second example discussed was from Johnston and Dark (1982). The subjects monitored dichotically presented words for any occurrence of a state's name. Sometimes they monitored a single ear and other times they monitored both ears. In addition, occasionally the primary task was interrupted by a free association measure (e.g. show "bark"; produce "dog" or "birch"). The critical measure was the degree to which the associations were influenced by prime

words in the prior text (e.g. "growl"). These prime words were expected to bias the associations given in the occasional test if they were semantically processed. There were three conditions: In the the selective attention condition with the attended prime, the association followed the prime 69%; in the unattended condition, it was 52% (50% was chance); and in the divided attention cognition, it was 59%. Thus, priming seems to be influenced by the filtering task. In this case, however, there is no significant priming for the unattended stimulus. For a strong late selection theory, equal priming would be expected in all three conditions. Thus this kind of result rules out completely unselective processing such as suggested by a strong late selection theory.

This is a particularly nice study because it provides controls that demonstrate how much semantic analysis there was in the selective attention and divided attention conditions as well as the unattended condition. Such controls are almost totally lacking in this literature. The technique used here to measuring the degree of selection failure can be profitably employed with some of the other indirect tasks such as the flanker task.

Pashler's four conclusions

Consider Pashler's four conclusions which are paraphrased here:

1. When people try to focus on certain stimuli and ignore others, they generally notice and report only relatively gross physical properties of rejected stimuli.

2. Under conditions that allow effective selections it is often possible to show that some semantic analysis of rejected stimuli still occurs on some trials.

3. When favorable conditions for attentional selectivity are provided, evidence of unselective processing is unconvincing.

4. Increasing the number of rejected stimuli reduces and even eliminates their effects, as indicated by the indirect measures.

We briefly discussed several issues: How representative are these conclusions of the cited studies? Are they too conservative ("lame") to be interesting? What studies are the best examples of evidence for each?

What would my Grandmother think of these conclusions?

David Brainard suggested we subject Pashler's conclusions to "the Grandmother Test". Namely, would these conclusions surprise your grandmother? There was general agreement that none of the four conclusions would pass the Grandmother Test. Each was carefully worded to not seem surprising. On the other hand, several of the original demonstrations of semantic processing of unattended information would probably surprise Grandma, especially the Stroop. Thus, there was a feeling that Conclusion 2 did not capture the impact of a phenomenon such as Stroop.

Is there Unconscious Semantic Priming?

Richard Abrams described his studies with Tony Greenwald on unconscious semantic priming. Their study took semantic priming from flanking stimuli as their departure point. Flanker effects have been observed when irrelevant words are presented above and below the central target word. This is taken as a failure of selection. Broadbent and Gathercole showed that this result was eliminated when novel words were used on each trial. In Abrams and Greenwald's experiment, the subject's response was a function of the semantic category of the word in the middle. If the flanking words belong to the different category than the target, reaction time increased. They constructed primes that shared syllables with the target words. These primes were effective. In fact they constructed a case where priming by parts of the word had an opposite meaning to the whole word and the parts were effective and not the whole. This effect suggests that partial cues may be mediating the priming. This is similar to results showing a few letters can produce some of the Stroop effect (Singer, Lappin and Moore, 1975). What is mediating this effect? By repeatedly being exposed to a set of words during the experiment, a subject may improve his performance by using partial cues to the particular words in the experiment. Thus, this experiment shows that evidence that appeared to favor semantic analysis of unattended words is actually much more specific and may be motivated by a subtle strategy to improve performance on the primary task

What about other kinds of perceptual analysis besides semantic?

Pashler's chapter described filtering tasks with a variety of selection criteria and a primary task that almost always involved identification of letters or words. Hui-Lin Chien asked whether the effects of selection were independent of the primary task. (This questions could also be asked for the

ease of selection.) Consider a display of multicolored geometric figures where one has to filter on one attribute (e.g. shape, size, or color). Then vary the primary task to be a report (or discrimination) of one of the other attributes. What would one remember about the irrelevant attributes or irrelevant stimuli in this task? Would one expect a flanker effect in such a task? Would it depend on the match between the selection criterion and the primary task? This is another approach to the issue of the fate of rejected stimuli.

Is location special for visual selection?

Pashler reviews the evidence for location playing a special role in selection relative to other attributes such as color. One line of thinking is related to what Graham calls "multiply selective" channels and what Watson and others call "labeled lines". The idea is that some internal processes are selective to both a particular feature such as orientation or color and spatial location. If all such processes always paired a feature with location, then location would be special in some sense. This analysis is supported be several experiments relating location and detection performance. The question is whether the mediating processes are multiply selective to other pairs of attributes besides location? Pashler also reviews evidence for location being special based on the pattern of errors. For example, intrusions for nearby locations compared to intrusions from similar colors. Such arguments, however, may be subject to the same issues of matching discriminability as discussed above.

Erin Harley pointed out that that location is treated differently than other attributes in most experiments. Typically each stimulus has a unique location while it is unusual for each stimulus to have a unique color or orientation. Proper comparison experiments need to control for this bias to distribute stimuli across space in a way different than they are distributed across other feature values. This thinking also leads to experiments where multiple objects are presented at the same location.

Does one select by location or by object? We briefly discussed examples of tasks using overlapping stimuli and superimposed video sequences where location is the same for attended and rejected stimuli. These filtering tasks are not easily accounted for a selection mechanism based on location because the overlapping objects share the same location. One must represent the objects separately before selection can be effective.

John Palmer tried to relate these ideas to the hypothesis that selection

may occur at many places in the sensory system. The idea is that there is a hierarchy of processing that extracts more complex combinations of attributes at higher levels of the hierarchy. Perhaps filtering can occur at each step in the hierarchy. A nice feature of this hypothesis is to only filter out irrelevant information when it will interfere with a specific process (cf. Moran and Desimone, 1985; Desimone and Duncan, 1995).

What is the relation between eye movements and visual attention?

Pashler describes the "mandatory shift hypothesis": A shift in attention is mandatory before the execution of a saccade. Pashler makes a strong case for it, and the experiment described in the chapter is pretty convincing (Shepherd, Findlay, and Hockey, 1986). The discussion also introduced some details from the Kowler studies on the subject. Several examples were introduced to try and make plausible the alternative hypothesis of being able to move one's eyes independently of where one was paying attention. Some examples came from the multiple perceptual demands during driving and walking such as attending to the road ahead while checking dashboard instruments. As another example, Jim Phillips proposed the inspection of a large object. Perhaps one can let the eyes wonder on an object while attending to a visual discrimination elsewhere? Despite the evidence, there was a general skepticism of the mandatory shift hypothesis.

We then briefly discussed alternative interpretations of these results. To what extent are these studies evidence for a common selection process versus some kind of capacity limitation between competing perceptual and motor processes? It would be useful to generate some alternative hypothesis along these lines.

Survey of Attention
Questions for the chapter on divided attention
John Palmer
21 January 1999

Study Questions

1. How does Pashler define perceptual capacity limitations?

2. How does the interpretation of searching a brief display depend on assumptions about noise?

3. How does the interpretation of simultaneous vs. successive presentation depend on assumptions about masking?

4. What is the double detection problem (also called the two-target problem)?

5. How similar are results in vision and audition?

6. What are Pashler's conclusions? What stimuli allow unlimited vs. limited capacity perception? (Hint: features and easy disc. letters vs. hard disc. letters and words)

7. Pashler's arguments depend on processes beyond perception introducing other kinds of capacity limits. List all of the other limits he considers in this chapter? (Hint: sensory, memory, decision, and double detection)

Discussion Questions

1. Pashler uses the term perceptual capacity to refer to the capacity of the perceptual system. In contrast, Broadbent refers to the capacity of the entire system and Townsend refers to the capacity of a single component process. What are the consequences of these different usages?

2. Pashler concludes that report experiments say little about perceptual capacity. Is there still a way to test theories of divided attention using report?

3. In this chapter, Pashler does not emphasize the distinction between dual tasks with separate responses for each stimulus and monitoring tasks with a single common response (but see Chapter 6). He would agree, however, that dual tasks differ from other monitoring tasks in avoiding decision noise at the cost of introducing possible interference between later processes such as memory and response. Do we need a separate section considering the interpretation of dual tasks?

4. The results of diverse studies are summarized by saying that there are no perceptual capacity limits for simple discriminations but there are for complex discriminations. This leads to asking what makes something complex. What are some alternative hypotheses?

The Counterpoint

Pashler describes the potential for noise to affect speeded search but he still concludes that large set-size effects rule out unlimited capacity perceptual processing. In contrast, Palmer and McLean (1995) argue that unlimited capacity models can produce any magnitude of set-size effect. Thus, large set-size effects cannot be used to rule out unlimited capacity. Is Pashler's interpretation a mistake or are there ways to salvage his argument?

Survey of Attention
Synopsis of Discussion
Week 3: Divided Attention
Hui-Lin Chien and John Palmer
3 February 1999

Primary Reading

Pashler, H.E. (1998). The Psychology of Attention. MIT Press. Chapter 3.

Background Reading

Shiffrin, R. M. & Gardner, G. T. (1972). Visual processing capacity and attentional control. Journal of Experimental Psychology, 93, 72-82.

Kleiss, J. A. & Lane, D. M. (1986). Locus and persistence of capacity limitations in visual information processing. Journal of Experimental Psychology: Human Perception and Performance, 12, 200-210.

Counterpoint Reading

Palmer, J. & McLean, J. (1995). Imperfect, independent, parallel search yields large set-size effects. Talk presented at the meeting of the Society of Mathematical Psychology, Irvine, CA. Transcript available.

What is the functional role of iconic memory?

Tony Greenwald asked what everyone thought of Pashler's comments about the functional role for iconic memory. Pashler repeated the argument that conditions under which iconic memory is observed in the laboratory are not typically encountered in natural environments. Tony Greenwald argued that iconic memory is necessary for activities such as watching television and movies because of the flicker in these displays. Further discussion raised the distinction between the temporal processes that mediate temporal integration and their description as a memory system. Clearly, temporal integration mediates some perceptual tasks, but is this equivalent to saying that iconic memory mediates those tasks?

How does the decision noise hypothesis account for accuracy search?

The accuracy of searching brief displays declines with increasing set size. On the face of it, this might be due to a capacity limitation in perception. Pashler cites the argument due to Shaw and Palmer and others that this may be due to the need to integrate information from more stimuli with the larger set size. Each stimulus give another chance for an error. Thus, the effect of set size is purely a statistical problem due to the increase of noise with the increase in set size.

Added after the seminar: This argument can be applied at either the time of decision as in Shaw's "independent decision" model or at time of perception as an integrative process as in Graham's "maximum or outputs" combination rule. These two theories make identical predictions for these kinds of tasks. Integrating information by taking the maximum of the relevant evidence yields the same result as making independent decisions on each piece of evidence. The point is that an early perceptual process version produces the same predictions as a late decision process account. Thus, this general hypothesis may be better termed the noise integration hypothesis rather than being labeled as specific to decision.

What is the main drawback of the decision noise account?

The decision noise hypothesis provides an alternative account of set-size effects that doesn't depend on capacity limitations in perception. A disadvantage of this approach is that several assumptions about the representation have to be made to obtain a specific prediction of the magnitude of the set-size effects. Thus, this theory is hard to test if one isn't committed to these assumptions. In contrast, the simultaneous vs. successive paradigm discussed next predicts no difference between conditions for the unlimited capacity hypothesis. Thus, it makes a specific prediction without as specific assumptions about the representation.

How is the simultaneous vs. successive paradigm interpreted?

Erickson and Spencer (1969) and Shiffrin and Gardner (1972) developed the simultaneous vs successive paradigm to test if perception has unlimited capacity. The rationale is that if capacity is limited, then subjects should be less accurate if all items have to be processed at the same time (simultaneous presentation) as compared to a condition where different items can be

processed at different time (successive presentation). They found that the difference between simultaneous and successive in performance was negligible for several conditions. This result was found for discriminations between letters and digits as well as specific letter pairs. It was also found for successive displays with a generous 0.5 seconds to switch attention. In contrast, a successive advantage is found for some complex letter discriminations and word tasks.

Jennifer McLean questioned whether this paradigm was without its own assumptions. Specifically, the physical displays differed in simultaneous and successive conditions. These differences could have produced differences in sensory processes (e.g. lateral masking, perceptual organization, texture cues?). Perhaps the changes in the displays minimized any simultaneous and successive difference? One way to resolve this problem is to always present two successive displays of 4 characters. Then compare the following two conditions: In the "simultaneous" condition, instruct subjects to only attend the first display. In the "successive" condition, instruct subjects to attend to two of the items in the first display and two of the items in the second display. This would allow the same comparison to be made with an identical set of displays (cf. Palmer, 1994).

What is the role of masking in the simultaneous vs successive paradigm?

Another potential concern with this paradigm is the role of the masks. If masks are treated the same way as characters, then the differences between the conditions is eliminated. This possibility was addressed in one of Shiffrin and Gardner's control experiments, but caution in necessary. A perhaps better comparison is the original Erickson and Spencer study that obtained essentially the same results without masking. The similarity of results with and without masking is reassuring. More generally, the effects of masks are assumed to be a simple interruption of previous processing. This may not be true and complicates predicting the size of the successive advantage with more complex stimuli (cf. Kleiss & Lane, 1986). Indeed, interpreting the cases with differences between simultaneous and successive presentation is the more troubling condition with this paradigm. These conditions have not been investigated both with and without masks.

What is the two-target effect?

The two-target effect was summarized as follows: Duncan (1980) used a

dual task with separate responses to separate groups of characters. When two targets were presented, accuracy declined relative to the one target condition. Duncan took this effect as evidence supporting the hypothesis that the capacity limitation occurs at a later stage. The process of detecting a target does not run up against capacity limitations, but processing a target does. The limitation must occur after detection because it pertains to targets and not distractors.

Can this conclusion be generalized?

Mary Czerwinski and Tony Greenwald questioned the generality of this analysis. For example, will it be easier for subject to simultaneously detect if one target is a light and the other is a tone? Tony cited a study that shows no capacity limitation when two targets are in different modalities (e.g., a light and a tone). Further studies by Bonnel and Hafter (1998) suggest that this depends on the details of the task. In these studies, there is no interference in detection but there is interference in identification.

Tony Greenwald also cited the McGurk effect (McGurk & MacDonald, 1976) as a relevant example. In this effect, a subject both listens to a phoneme and sees a video of a phoneme being pronounced. If inconsistent phonemes are presented, there is considerable interference with visual input often modifying what a subject reports hearing. Thus, this is a drastic case of representations not being independent. Tony pointed out that all of the analysis in this chapter depends on assumptions about what are the channels or representations that subjects use to base their judgments. For example, the individual letters in a search task are assumed to result in a set of corresponding independent perceptual representations.

John Palmer described a similar case in his search experiments where one of his subjects found a strategy to use the whole stimulus set as a pattern. The result was performance that improved with larger set sizes. The strategy was defeated when additional randomization of position made the patterns more unpredictable. The point of these examples was both the need to make these assumptions more explicit and to question the generality of this analysis.

Is there no perceptual mediation of the two-target effect?

Hui-Lin Chien pointed out that two-target effect might still have a perceptual component. For example, does the distance between the two

targets on the display affect the two-target effect? Such a distance effect would be consistent with some kind of perceptual limitation mediating the two-target effect. Alternatively, if the two-target effect is due to memory encoding, then one should see interactions with memory variables. There is little positive evidence one way or the other at this point.

What about other perceptual limitations?

Tony Greenwald asked about how to interpret restrictions on the spatial distribution of attention. If processing was unlimited, then why are there any constraints on selection. Imagine a 4 by 4 display with 16 letters. Can you report an arbitrary 4 letters out of this array? That seemed implausible. But others countered that perhaps you could report an arbitrary 4 items if they were a different color than the others. The thrust of the question is how to relate constraints in selection to the lack of constraints in perceptual processing capacity.

Indirect measures of perceptual processing capacity.

Tony Greenwald described an indirect measure of perceptual processing capacity. The idea was a variation on the unconscious priming experiments of his and his students. Primes are presented, masked, and then subjects are to make an affect judgment of a target word. Is this word positive or negative in emotional content? The primes affect this judgment even when they are very hard to detect. The new twist was to introduce uncertainty into the location of the prime and then cue its location. Interestingly, the amount of priming was increased with a location cue relative to no cue. Thus, does attention facilitate the processing of the prime word that can't be reported later?

John Palmer countered with a noise integration account of this effect. The design is similar to the multiple mask conditions of Shiu and Pashler (1994) to be discussed next week. The idea is that without the cue, the effect of the prime is weakened by including the noise from the areas of the blank screen that were followed by a mask. With the cue, noise from these locations could be excluded. Thus, this indirect tasks has the same issues of interpretation of a direct measure such as used by Shiu and Pashler.

What stimuli allow unlimited vs. limited perceptual capacity?

Pashler summarizes the results of many studies in terms of simple stimuli

being processing with no limitations on perceptual capacity while complex stimuli do have such limits. Simple stimuli include features such as color and orientation, and some letter discriminations. Complex stimuli include words, some letter discriminations, and relations between different objects. Discussing this summary led to a few comments about testing word stimuli and analyzing compound words such as "lipstick".

Note added after seminar: A problematic issue is how to interpret the mixed results found with letters. Does it mean that some letter pairs and tasks require limited capacity processing and others don't? Or, does it mean that some tasks can be accomplished with only partial identification. Perhaps full identification always requires limited capacity processing? From this point of view, even the digit vs letter discriminations may be not revealing the true nature of letter identification.

What does Palmer and McLean's work tell us about the set-size effect?

This discussion begins from the counterpoint article (Palmer & McLean, 1995). This article focused on the set-size effect in response time which is widely used to distinguish unlimited capacity, parallel processes from limited capacity, parallel or serial processing. Palmer and McLean argue that an unlimited capacity, parallel model can account for set-size effects of any magnitude.

Jennifer McLean described the main point as not being able to infer a specific (e.g. serial) model from observing a large set-size effect. The paper provides a counterexample in showing how large set-size effects can result from unlimited capacity, parallel model. There was then some discussion of the relation between this work and Townsend's.

Tony Greenwald questioned what this result really tells us. Jennifer replied that their paper was intended primarily to clarify the relations among alternative models rather than resolve the issue. In addition, the paper also argues that the data available are inconsistent with a particular limited-capacity, parallel model.

Tony questioned whether this kind of theory was useful compared to simply stating the empirical generalizations. John Palmer argued that both were useful. He agreed with Tony that generalizations such as the huge effect of discriminability on set-size effects were important to emphasize and have

sometimes been lost in the debate about parallel vs serial models. On the other hand, John also argued that theory was necessary to get beyond a list of unrelated observations.

What are the common alternative definitions of capacity?

This discussion began with a story describing a phone conversation with Barry Kantowitz. Barry argued that Pashler used the concept of capacity incorrectly right from the start of the chapter. The argument is based on a conflict between Pashler's usage and Townsend's usage. Townsend makes a theoretical dichotomy that distinguishes between system architecture (serial vs parallel) and the capacity of individual processes (limited vs unlimited). In contrast, Pashler describes a prime example of a limited capacity as being serial processing. Thus Pashler makes serial processing an example of limited capacity and Townsend doesn't.

One way to understand the issue is to begin from the four kinds of processing in Townsend's dichotomy: unlimited-capacity parallel, limited-capacity parallel, unlimited-capacity serial and limited-capacity serial. Here, capacity refers to the individual process not perception as a whole. Pashler's view is to label one of these cases, unlimited-capacity parallel as unlimited perceptual capacity. This case is also called the independent channels model. This approach emphasizes that only one of the four models shows unlimited capacity at the level of the entire perceptual process.

In summary, the difference in usage can be understood as applying capacity to either individual component processes or to the entire stage of processing (e.g. perception). Indeed, Broadbent uses the term in a third way when he talks about the capacity of the entire information processing system (stimulus to response). Thus, there are three different usages of the term in common use.

Pashler's chapter consistently uses the term with respect to perceptual capacity. The only point that is unclear is when he references Townsend's work in the beginning of the chapter. At that point, it would be useful to elaborate this distinction to avoid confusion with Townsend's use of the term. In the long run, I expect all three uses of the term will prove useful. But hopefully the terminology will improve as more authors recognize the multiple usages.

Survey of Attention
Week 4: Questions for the Chapter on Attentional Set
John Palmer
4 February 1999

Study Questions

1. What distinguishes set experiments from filtering and monitoring experiments?

2. What are the differences between capacity allocation and noise reduction?

3. For what simple attributes are there effects of set in detection?

4. What is the evidence for capacity allocation vs. noise reduction in the single-element, discrimination accuracy experiments? How about the speeded discrimination experiments?

5. What evidence is there for an effect of set on the perception of object identity?

Discussion Questions

1. I like the probe-signal paradigm. Might it be particularly good in producing large effects? What issues can it address? What issues is it not particularly good for?

2. Pashler finds very little evidence for an effect of set on object identity. This is inconsistent with the major role of top-down processing in some theories of object and scene perception. Are these theories wrong or have important phenomena been missed in these experiments?

3. Pashler argues that the most critical contrast of the capacity allocation and noise reduction hypotheses depends on the benefit of precues relative to an appropriate neutral condition. Why are benefits more critical than cost? Furthermore, the size of a benefit depends on using an appropriate neutral condition. What influences performance in the neutral condition?

4. For speeded discrimination, would a speed/accuracy tradeoff experiment resolve the differences in interpretation of the effects of set?

The Counterpoint

What can one learn by contrasting Shiu and Pashler (1994) against Luck et al. (1996)? What are the assumptions about representation and the assumed limits on performance? What experiments might further illuminate the conditions that produce effects of set? How can one resolve the results in these set experiments with the results of divided attention experiments such as accuracy visual search or the simultaneous/successive comparison?

Survey of Attention
Synopsis of Discussion
Week 4: Attentional Set
Erin Harley and John Palmer
12 February 1999

Primary Reading

Pashler, H.E. (1998). The Psychology of Attention. MIT Press. Chapter 4.

Background Reading

Shiu, L. & Pashler, H. (1994). Negligible effect of spatial precuing on identification of single digits. Journal of Experimental Psychology: Human Perception and Performance, 20, 1037-1054.

Counterpoint Reading

Luck, S. J., Hillyard, S. A., Mouloua, M. & Hawkins, H. L. (1996). Mechanisms of visual-spatial attention: Resource allocation or uncertainty reduction. Journal of Experimental Psychology: Human Perception and Performance, 22, 725-737.

What is Pashler's definition of a set effect?

Pashler gives a broad definition of attentional set on page 167. He defines the effect as, "whether having advance information about a stimulus can help one perceive that stimulus more effectively." Someone asked if this is similar to priming? This definition may include priming, but unlike priming, most set studies explicitly direct subjects to attend prior to the task. Perhaps set can be defined simply as predisposing someone, giving them advance information, framing a situation or giving it context. An example of this would be telling someone they will see a word related to the word "bank" and then presenting them with either "money" or "groceries." Response time will be faster after the presentation of the word "money." But is this priming really attentional?

John Palmer suggested that the idea of attentional set is much like

divided attention and asked why Pashler separated attentional set out into its own chapter? Perhaps the last chapter was already too long? A better reason may be that the set experiments deal with single stimuli while the previous chapter deals with multiple stimuli. However, with low threshold stimuli, a blank screen can be confused with a stimulus which weakens this distinction.

Are there good examples of set in ordinary experiences?

Pashler introduced set in part from the early writings of Lucretius who claimed that things are not seen sharply "save those for which the mind has prepared itself." Richard Abrams pointed out that Lucretius was probably not thinking about location cueing experiments when he made that statement. Barry Kantowitz gave the real world example of driving along highway 101 and seeing a "Beware of Elk" sign. He suggested that the sign primes him for seeing an elk, and probably improves his braking performance.

How important are these studies?

Barry went on to cite Broadbent as arguing that scientists have an obligation to study important things. It seemed to some that perhaps the real thing being studied in these experiments are the physical properties of masks, not real world set effects. Is the attention research too hung up on theories? Which is more important, the theory or the results? Barry challenged the group: "who cares about the theory!" He gave an example of an application in masking of turn signals by headlights. He concluded that the value of attention to real world applications do not come through in Pashler's chapter.

Shiu and Pashler example study

We next turned to the substance of the chapter. Andrea Hartzler gave a summary of Shiu and Pashler (1994). The experiments were done in response to Henderson (1991) and others who found cueing effects both with single and multiple stimuli in a display. Shiu and Pashler suggested that using multiple masks in Henderson's study generated the observed effect. To test this idea, they compared single mask displays to multiple mask displays. Valid, Invalid and Neutral cues were all tested. Shiu and Pashler found cueing effects for the multiple mask trials and no cueing effect for single mask trials. Their conclusion was that the masks used by Henderson created ambiguity in where the stimulus was displayed. Thus the decision was based on more than one location and the other locations added noise that caused the subjects to make additional errors.

Thus, their study provided support for the noise reduction interpretation of these effects.

Alternative cueing experiments

The use of partially valid cues introduces a complexity in interpreting these experiments. Compare them to the case in which cues are 100% valid or are neutral. In the 100% valid condition, subjects should use information from only the cued location. Performance relative to the neutral control can be predicted using the noise reduction models worked out for divided attention. The situation is more complicated with partially valid cues. Consider a single channel model in which it is possible to solely focus on the cued position in a display. Given such a model, subjects will always miss the target with invalid cues. Furthermore, they will miss with or without multiple masks. Now if subjects know that they cannot trust the cues, they probably will monitor other locations in the display. The cueing effect would arise from the relative weight given the different locations, but there is nothing in the instructions that specifies the appropriate weight. Thus, one cannot predict the specific performance expected in this experiment. The 100% valid cue experiment is more constrained because the appropriate weights are clear from the instructions.

Alternative versions of the noise reduction hypothesis

We discussed the alternative hypotheses laid out by the Figure at the beginning of the chapter. Someone asked about the distinction between an early selection process and a late "selection" using weights in decision. Can these be distinguished? One result supporting early selection from prior chapters was that motion adaptation (e.g. the waterfall illusion) is stronger when one attends to the motion (e.g. Lankheet & Verstraten, 1995).

We went on to discuss the detective analogy used in the chapter to explain the idea of noise reduction. Some found the analogy confusing. To use the analogy to understand noise reduction, you have to accept that more detectives is not necessarily better.

Is there an alternative theory that avoids the effect of noise reduction?

These models all assume a signal detection view of detection and discrimination, but what if we embrace a high threshold model? A high

threshold model assumes that no distractor can ever put you in a detect state. You may miss the target, but you would never confuse a non-target for a target. This would imply no false alarms other than those due to guessing. If the high threshold model is accurate, then one expects no effect of multiple masks. The high threshold model is an older idea that was common in early psychophysics and remains in some response time models. For example, all of Townsend's serial-parallel equivalences are on the high threshold model.

Why would one ever assume a high threshold model? On the plus side, it is a far simpler model than the signal detection theory. The math is drastically simpler and one can make many specific predictions that are hard to characterize for other models. The negative side is that it is too simple; many of its predictions are known to fail. Barry Kantowitz felt it should be a reminder to look for purity first. If a simple model can explain for enough of the phenomena, than why look further for a more complex model which does not make clear predictions? He contended that sometimes we are forced to go further reluctantly. The main point for the current discussion is to understand that the only alternative to some kind of noise reduction model is to adapt the high threshold model.

Relations between the probe signal paradigm and partially valid cueing

In a typical probe signal experiment, subjects are asked to listen for a tone and make a detection response. Suppose subjects are trained with a tone of 1000 hz. During the actual trials 500 hz tones were occasionally slipped in between the 1000 hz tones. Subjects do poorly at detecting these 500 hz tones when they were listening for the 1000 hz tones. The account of these effects is that subjects are attending to a single frequency channel and ignoring other channels. This produces very large effects similar to the fall off in a single channels sensitivity to frequency. In summary, inducing subjects to adapt a single channel strategy may be a way to generate particularly large attention effects.

Someone asked whether attention mediates this effect? Pigeons can be taught to respond to the presence of a light, and then do poorly when the light is changed. With pigeons this is explained as stimulus generalization and discrimination. Maybe the effect in people is a similar effect of learning? The arguments for attention are primarily that careful experiments can alternate between mixed frequency blocks and blocks where most of the tones are a single expected frequency. The variation is signaled by instruction to the

subject. In addition, the fact that people can attend to disparate frequencies and ignore the intermediate frequencies supports a hypothesis with more than a simple stimulus generalization.

The connection to the Shiu and Pashler study is that the set they are establishing may encourage subjects to ignore the uncued location even though they could process it. This strategy may contribute to differences between valid and invalid cues; but, it will not contribute to differences between valid and neutral cues. Thus, the benefits of cueing (as they are called), are not susceptible to this explanation. For this reason, the occurrence of benefits is particularly strong evidence for a limited capacity interpretation.

What is the appropriate neutral cueing condition?

What is a good neutral cue condition? One version is to cue every location, and another is to cue no locations. Shiu and Pashler used the latter in their experiments with single versus multiple mask conditions for single stimulus displays. The potential problem with this version is that the absence of a cue may give the subject less warning of when the target is going to be displayed on the screen. Thus, the cue type may be confounded with a change in temporal uncertainty.

Luck et al. were able to demonstrate the cueing effects for single stimuli displays by increasing the stimulus onset asynchrony (SOA) between the cue and the stimulus. They argued that Shiu and Pashler did not give subjects enough time to shift their attention after the cue. Shiu and Pashler's SOA was only 100 ms, Luck increased this to 250 ms. Why stop there? If you make sure subjects are not moving their eyes why not give them a second between the cue and the stimulus onset (cf. Palmer, 1994)? Luck also used four different cues in his experiments to see if the type of cue affected the magnitude of the effect. When cues were too close to the stimulus location, subjects did worse with a cue compared to no cue. Luck reasoned that these cues were masking the stimuli. Perhaps an auditory cue would solve this problem. Alternatively, a blocked design with verbal instructions could be given instead of visual cues. In such an experiment subjects are told, "the target will always be in the upper right corner of the display." This could be compared to blocks in which they are told the stimulus will appear in multiple possible locations. Such blocked conditions require particular attention to the issue of maintaining appropriate eye fixation.

The difference in performance between valid and neutral cueing conditions is small at best. What if any conclusions can we draw from such small differences? Perhaps occasional lapses in attention could give us these small effects. Since performance is averaged across trials we do not know what subjects are doing on a trial to trial basis. Some argued that pursuing this experiment work is a waste of time because the effects are so tiny. Others argued that the goal of these experiments is to take a weak argument for the capacity allocation model and make it even weaker.

Response time studies of single stimulus location cueing

In simple response time studies, subjects are asked to press a button as soon as they see the stimuli. Suppose there are four possible positions for the stimuli to appear. Subjects are faster when a stimulus appears in a cued location compared to an uncued location. One account of this result is that the cue facilitates perception at the cued location. An alternative account is that the cue lowers the response threshold for the cued position and raises the threshold for uncued positions. Such decision criteria can account for the the change in response time. Simple response time experiments don't have an obvious way to discriminate these possibilities.

With choice response time, subjects must identify which of two or more stimuli they have seen, for example a "5" or a "6" as in Shiu and Pashler. Typical results again show that subjects are faster when a stimulus appears at a cued location compared to an uncued location, but again there are two interpretations. Either perception is facilitated or the decision criterion is adjusted. However, now the criterion model predicts an increase in errors in exchange for the decrease in speed. Indeed, Pashler cites some evidence that this is what happens. A more complete analysis of this possibility would measure the entire speed/accuracy tradeoff function. With the criterion hypothesis, the speed/accuracy function does not change, the criterion shifts simply move performance along the curve. Alternatively, if perception is modified by the cue, then the entire curve shifts when the stimulus is cued. Such an experiment has not yet been done.

Object identity cueing experiments

The cueing experiments above all involved cues that were orthogonal to the attribute being judged: location cues for letter discrimination, frequency cues for tone detection, etc. Here the cues provide direct information about

the identity of the objects being identified. Such manipulations often affect the bias for one response or the other. To control for this, Pashler emphasized comparisons where the cue is given both before and after stimulus presentation or is given just after stimulus presentation. Object identity cues often show no differences between the before-and-after and after conditions. However, Pashler does describe one case where they do have substantial effects. Subjects were asked to identify a pictured object. They were given response choices either before or after they had viewed the object. When discriminating between large categories such as "shoes", "elephants", "telephones" and "cars", subject performance was no better in the before-and-after than the after condition. However, when subjects had to discriminate between similar objects such as four different styles of shoes, they did better in the before-and-after than the after condition. Pashler suggests the cue helps when it allows the subject to encode different kinds of information than they would encode usually. For example, one doesn't usually remember details about shoe styles, but can if relevant to the task.

What can we say about the role of attention in object recognition? Maybe attention is only needed in object recognition when there are a multiplicity of meanings or when the meaning of an object is ambiguous. Certainly knowing the context of speech before reading or listening to it aids in understanding.

The seminar ended with a spontaneous effort to design a set experiment using ambiguous letter/digit stimuli. Consider a character that can be interpreted as an "A" when presented in the word "CAT", but is interpreted as an "H" when presented in the word "THE". Could one cue these alternate interpretations? Would this cue be more effective in the before than after condition? Suppose the cues were either "A" or "H" and were presented either before or after the ambiguous letter. The response might be to identify the letter. Would such a cue have more influence in the before than the after condition? Perhaps top-down influences would be more potent given ambiguous bottom-up information? On the other hand, it may be that top-down influences are not effective even in this case where they have their best shot.

Survey of Attention
Week 5, Chapter 5: Questions about the Theory of Attention and Perception
John Palmer
11 February 1999

Study Questions

1. What are Pashler's 12 main empirical generalizations?

2. Which are critical to rejecting early selection theory?

3. Which are critical to rejecting late selection theory?

4. Does the controlled parallel theory predict all of the generalizations?

5. What are the differences between the exclusion and capacity allocation views?

6. What are the arguments for sharing vs switching? Particularly fast switching?

Discussion Questions

1. To what extent are Pashler's 12 main generalizations in dispute? What is needed to resolve each of these disputes?

2. How convincing is the evidence against early selection?

3. How convincing is the evidence against late selection?

4. How good is support for the controlled parallel model? What are alternatives?

5. Pashler argues that either selection or capacity can be considered central to attention. Do you agree? Do we have to pick one or the other?

6. Pashler suggests little role for preattentive processing. How can this view be reconciled with the typical two-stage search theory (e.g. Treisman or Wolfe)?

7. Pashler argues against fast switching of attention. Does this mean that all serial processing models in visual and memory search are fundamentally misguided?

The Counterpoint

How effective are the defenses of early selection by Luck et al. (1996) and of late selection by Duncan (1980)? Do they counter any of Pashler's arguments?

Survey of Attention
Synopsis of Discussion
Week 5: Theory of Attention and Perception
Richard Abrams and John Palmer
22 February 1999

Primary Reading

Pashler, H.E. (1998). The Psychology of Attention. MIT Press. Chapter 5.

Background Reading

Duncan, J. Ward, R. & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. Nature, 369, 313-315.

Counterpoint Reading

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. Psychological Review, 87, 272-300.

Luck et al., (1996) from last week.

Outline of discussion

Pashler's chapter first offers a critical recapitulation of earlier chapters, followed by a presentation of new material. John Palmer announced that today's meeting would follow the chapter in reverse order: new material, then recapitulation.

How long does it take to switch attention?

Hui-Lin (Sarina) Chien presented an experiment from Duncan, Ward and Shapiro (1994). In a dual task paradigm, they showed poorer performance when the 2nd task followed the first by up to 500 ms. They argued that this reflected the period of time that attention was committed to the first task and could not be allocated to the second task.

We briefly discussed similarities and differences between this paradigm and both the attentional blink and the PRP paradigm. First reactions to Duncan's results were that they stand as evidence against a fast-switching

capability. But questions were raised about the exact nature of the switching in the experiment. Anthony Greenwald expressed the view that not only was spatial location being switched, but also the instructions specific to the task (i.e. from a letter to a number discrimination, or vice versa). This concern was reduced for the 2nd experiment of the paper that used a single search task with a single response. A similar point was raised by Barry Kantowitz, who suggested that the switch involved a change of categories. Barry wondered whether a switch not involving category change (e.g., from a 2 vs. 5 to a 6 vs. 7 discrimination) would produce the same result. Again the 2nd experiment helped resolve this concern. Barry also asked how this related to the PRP paradigm. Perhaps the subjects were making implicit responses to each stimulus and those responses were causing the interference. Again, the second experiment made this interpretation less likely. To sum up, John asked the seminar whether Sternberg was wrong: do we accept Duncan's results as refutation of Sternberg's influential findings of rapid serial search. Anthony Greenwald answered with a resounding "Yes," a sentiment that the rest of the participants accepted without dispute.

How convincing are Pashler's empirical generalizations?

Each of Pashler's 12 generalizations was addressed in turn. (A copy of a handout listing these generalizations is at the end of this summary.) For the most part there were only minor objections to Pashler's first set of generalizations concerning selection and the fate of unattended information. Barry Kantowitz suggested that in the generalization claiming no direct memory for unattended stimuli (Point 1), the word "no" distorts the empirical evidence and should be replaced by "little." John pointed out, however, that many procedures involving unattended stimuli are vulnerable to "leakage" and that it is difficult to distinguish effects of truly unattended stimuli from effects of stimuli supposedly ignored but intermittently attended. Thus, one of Pashler's points is that it is hard to distinguish "no" from "little".

Does selection depend on discriminability alone?

Two other points were made regarding the first set of generalizations. While difficulty of selection clearly depends on discriminability, the critical aspect of discriminability may not be some physical dimension (Number 2). This qualification was reflected in John's putting the adjective "physical" in parentheses in his restatement of the generalization—"Selection depends on (physical) discriminability."

How complete is selection?

The next point concerns Pashler's generalization Number 3, that the semantic analysis of to-be-ignored stimuli has been revealed in many studies through indirect measures. John emphasized the fact that while the extent to which to-be-ignored stimuli are analyzed is still being debated (cf. Broadbent & Gathercole, 1990), the evidence is clearly against unselective analysis of to-be-ignored stimuli.

What complex categories introduce limited capacity perception?

The second set of generalizations concerning divided attention was accepted largely without objection. The exception was Number 5: parallel, unlimited capacity search is possible when targets belong to simple categories, e.g. letters vs. digits. It was suggested that the case is more complex than this generalization suggests, and that unlimited capacity may apply only to certain kinds of targets that belong to simple categories. When targets are confusable (e.g. letters and digits with similar features), sequential not parallel identification is found, despite the targets belonging to simple categories. This led to the observation by Barry Kantowitz that whether targets belong to simple categories or not may be less important than their discriminability at a lower, featural level. Richard Abrams pointed out that discriminations in search procedures involving masking may rely heavily on lower-level elements (the particular elements that escape masking). There was general agreement that "simple categories" does not adequately capture the characteristic of targets that enables unlimited-capacity parallel search. But no one had a solution to this inadequacy.

Are these generalizations useful?

How well do the generalizations as a whole summarize the book so far? Opinion was favorable. Mary Czerwinski expressed her relief that Pashler was finally giving a mixed model serious consideration. The mixed model that Pashler eventually advocates was approached via the question of how well other models—early and late—accommodate the 12 generalizations. Table 5.1 in Pashler (p. 220), which indicates whether each model is consistent or inconsistent with each generalization, served as the basis for this discussion.

How useful are discussions of early and late selection theories?

Is it meaningful to evaluate empirical evidence (as in Table 5.1) from the standpoint of theories like late- and early- selection, theories that are artificial, unrealistically, and extreme? John advocated that it is useful to have models—however unrealistic—that can be rejected on some grounds. One must narrow the field of available options. Barry Kantowitz argued that at the very least we should evaluate not some abstract late-selection model but instead, say, Duncan's 1980 model, and not some abstract early-selection model but perhaps Broadbent's 1958 model. There was agreement that moving away from vaguely stated models towards a specific model associated with a particular author, paper, and historical moment was a good idea. Furthermore, the labels for the models which suggest general properties of selection need to be replaced with labels that make explicit that an entire theory is being tested rather than the nature of selection alone.

What evidence rejects the late selection model (e.g. Duncan,1980)?

Some of the points used to reject the late selection model can be questioned. First consider Point 3 and Point 11: Selection affects indirect memory for the unattended, and more stimuli reduce this indirect memory. These generalizations are a problem for late selection theory only if one assumes that processes after selection cannot further increase indirect memory. Second, consider Point 6: Limited capacity for complex categories such as words. This is a problem for "late" but not for the idea of a filter in general. Third, consider Point 12: Capacity limits apply to attended stimuli. Why this is a problem for a filter theory? In sum, the arguments need to be made more explicit.

What evidence rejects the early selection model (e.g. Broadbent, 1958)?

Here the sole argument is Point 5: Unlimited capacity for simple categories such as letters vs digits. This point is not based on as convincing a set of data as desired. In particular, some letter discrimination experiments do show limited capacity. Perhaps the experiments with unlimited capacity are achieved by using partial identification based on simple features? More generally, at most this point rejects the "early" aspect of the model and not a filter theory in general.

The controlled parallel model

Discussion then turned to Pashler's proposed controlled parallel model. Barry immediately questioned whether there was a substantive difference between this and Broadbent's 1958 model. John held that there was a difference in the separate filtering and limited-capacity mechanisms. This led to animated debate over the exact interpretation that should be given to Pashler's somewhat whimsical depiction of the controlled parallel scheme in Figure 5.1. Particular uncertainty arose over what Pashler intended in his depiction of stimuli that had successfully passed through the filtering mechanism—is allocation of resources to be equated with semantic analysis as the figure might suggest?

To clarify the distinction between Pashler's and Broadbent's models, Tony suggested thinking of Broadbent's model in terms of a radio metaphor: radio waves are analogous to incoming sensory information; the tuner that selects only a single frequency (channel) is analogous to Broadbent's filter. In contrast to a single channel being selected by the filter in Broadbent's model, Pashler's model (in Figure 5.1) apparently allows multiple channels to be admitted by the filter.

John, hoping for further clarification, then sought to place Pashler's model in a framework of different processing stages: early unselective sensory processing; selective, unlimited-capacity perceptual processing; and, selective limited-capacity processing in memory. Thus, some of the processing after the filter is unlimited capacity and some is limited capacity. Barry asked if such a model has two filters and John acknowledged it probably would.

Barry argued that while this interpretation of the model might be distinct from Broadbent's 1958 model, it was essentially the same as Broadbent's later (1971) model. Here John agreed that they were similar.

Are there precursors to the controlled parallel model?

Tony found it surprising that Posner and Boies (1971) weren't mentioned by Pashler in the context of his model. Posner and Boies hadn't developed an elaborated theory, but they framed their discussion in terms of exclusion and capacity as separate functions—the two functions that are the basis of the controlled parallel model. John responded that these two concepts aren't new, that they predated Posner and Boies and have been absorbed into the general debate about attention. Thus it's not so surprising that Posner and Boies weren't specifically mentioned. John added that one of the important questions

that should guide modeling is what is the least perturbation of the filter model necessary to accommodate the empirical findings.

Can one reject the general filter model?

In discussions above, the evidence was convincing against early and late selection models, but was not convincing against the idea of filter models in general. In such a general model, the filter occurs at some point in the processing stream. Processing before it is unselective and unlimited capacity and processing after it is selective and limited capacity. Is there evidence that allows one to reject this the general idea of a filter? In particular, it needs to demonstrate that selection is independent of limited capacity. Such evidence is not obvious in Pashler's chapter. One source of such evidence may be based on the the attentional effects on motion adaptation (e.g. Lankheet & Verstraten, 1995). This effect shows selection on some aspects of motion processing. Can one then show unlimited capacity for the same motion processing? For example, search for speed increments has been shown to be compatible with unlimited capacity (Verghese & Stone, 1995). Thus, the idea is that selection limits adaptation but does not require limited capacity processing for successive motion processing. Does this rule out a general filter model?

How complex can the filter be?

More debate ensued about the possible roles of a filter. Tony described recent work by Sperling (Lu & Sperling, 1995, Science, 377, 237-239) that suggests a filter operating early in processing. In this work, stimuli are presented that can produce the impression of apparent motion in one of two possible directions. The direction of apparent motion depends on which combination of stimulus features are attended to (e.g., a particular orientation and a particular depth). Everyone agreed that this impressive result is consistent with an early filter for at least certain stimulus characteristics. However, Sperling himself doesn't interpret the finding in terms of a general filter mechanism, but in terms of specific other processes such as "setting up a salience map".

How is the complexity of the filter related to capacity limits?

Tony proposed the idea of linking capacity and filter by viewing capacity as the effort to construct and implement the filter required in any specific situation. As example, Tony once again used a radio analogy. Suppose you

want a car radio to select information about traffic conditions from all the information it receives. One way of doing this would be to search for the word "traffic" and to allow only information associated with that word to pass to the next stage of processing. According to Tony's proposal, the capacity of the system would be equated with the effort and resources involved in constructing and implementing this filtering mechanism.

In discussing this proposal, the following prediction was made. The more complex the filter, the more likely performance would be limited capacity. Thus, one might expect that searching for an increase in speed in a display made up of multiple 3rd order motion patches would show limited capacity relative to the same experiment with 1st order motion.

Chapter 5 Handout: List of empirical generalizations
John Palmer
11 February 1999

Selection and the Unattended

1. No direct memory for unattended stimuli

2. Selection depends on (physical) discriminability

3. Selection affects indirect memory for unattended

11. More stimuli reduce this indirect memory

Divided Attention

4. Unlimited capacity for features

5. Unlimited capacity for simple categories (letter vs digit)

6. Limited capacity for complex categories (words)

7. Limited capacity for multiple targets

12. Capacity limits apply to attended stimuli

Attentional Set

8. Orthogonal cuing benefits in noise only

9. Identity cuing affects bias only

10. ... except for long displays and ...

Note: Numbers and substance from Pashler. Organization and text from Palmer.

Survey of Attention
Week 6, Chapter 6: Questions about Attention in Sensorimotor Tasks
John Palmer
18 February 1999

Study Questions

1. What are AOC functions? What results indicate independence? Interference?

2. What is the PRP effect? How can one interpret graphs such as Figure 6.2.

3. What are Pashler's 4 predictions about interactions in the PRP effect (p 279).

4. What is the evidence against the graded capacity sharing hypothesis?

5. What is the evidence against the crosstalk hypothesis?

6. How are central capacity and perceptual capacity related?

7. Under what conditions does the PRP effect seem to disappear (p 274)?

Discussion Questions

1. What empirical generalizations summarize the research in this chapter?

2. What might temporal uncertainty and preparation contribute to the PRP effect? What might response production interference contribute to the PRP effect?

3. Pashler argues for distinct central and perceptual capacity. Are you convinced?

4. How does Pashler suggest one might generalize from the simple tasks of the PRP to more complex and realistic tasks such as speaking and driving?

5. How might one further specify the response selection hypothesis? How could one predict the magnitude of the PRP effect?

The Counterpoint

Consider Meyer and Kieras' (1992; 1997) account of the PRP effect. How do they differ from Pashler in their empirical generalizations? Interpretations? Theory? Goals?

Survey of Attention
Synopsis of Discussion
Week 6: Attention and Sensorimotor Tasks
Andrea Hartzler and John Palmer
24 February 1999

Primary Reading

Pashler, H.E. (1998). The Psychology of Attention. MIT Press. Chapter 6.

Background Reading

Pashler, H.E. (1990). Do Response Modality Effects Support Multiprocessor Models of Divided Attention. Journal of Experimental Psychology: Human Perception and Performance,16, 826-842.

Counterpoint Reading

Meyer D.E. and Kieras, D.E. (1992). The PRP Effect: Central Bottleneck, Perceptual-Motor Limitations, or Task Strategies? Talk presented at the meeting of the Psychonomic Society, St. Louis.

Meyer D.E. and Kieras, D.E. (1997). The PRP Effect: Central Bottleneck, Perceptual-Motor Limitations, or Task Strategies? Psychological Review, 104, 3-65, and 749-791.

What is the PRP paradigm?

Erin Hartley introduced the psychological refractory period (PRP) paradigm using Experiment 2 from Pashler (1990). The PRP paradigm entails sequential presentation of two stimuli where subjects make two responses, one to a first stimulus and one to a second stimulus. Response time to each response is measured as a function of the time between the stimulus onsets (SOA). The PRP effect is the increased response time to the second stimulus at short SOA. Often SOA is reported to have no effect on the first response time.

Pashler investigated how the PRP is affected by uncertainty in the order of the two tasks and the use of different response modalities. In this

experiment, three letters were mapped to keyboard responses and two tones were mapped to vocal responses. Pashler found the usual PRP effect when the stimuli were in a known order and a larger effect with unknown order. The magnitude of this increase was dependent on the choice of response modalities. It was smaller when manual and vocal responses were combined compared to manual and manual.

Might a speed/accuracy tradeoff contribute to the PRP?

Barry Kantowitz argued that a speed/accuracy tradeoff contributes to the typical form of the PRP effect. The example study did not have a very complete description of the accuracy data. Specifically, while no effect of SOA on accuracy was reported, it wasn't clear how sensitive was the measurement of accuracy. With high accuracy, the variability of estimating accuracy can be large. Moreover, when accuracy is high, a small change in accuracy can have a large effect on response time. This combination makes it difficult to determine if there is a small effect on accuracy that might be hiding a larger effect on response time. Barry briefly described previous work of his (Knight & Kantowitz, 1974, 1976) that observed effects on the first response as well as the second when both accuracy and time were carefully measured. In particular, under speed stress conditions, he found that there was an effect on accuracy for the short SOAs for the first response. With an accuracy stress, this turned into an effect on the response time instead. As a result, Barry was not convinced that the effects were limited to the second response without a more complete accuracy measurement. This discussion continued with brief consideration of the appropriate controls, introduction of temporal uncertainty, and variations in the paradigm such as the "Partial information Paradigm" (Kantowitz & Sanders, 1972).

John Palmer concluded this discussion arguing Pashler's side. He argued that the magnitude of effect due to the speed/accuracy tradeoff was probably much smaller than the main PRP effect on the second response. Indeed, Pashler's experiment was focused on measuring the leading edge of the effect at short SOAs and then a control at a long SOA. His experiment was not intended to measure the details of the effect of SOA over time or speed/accuracy issues. But this argument did not convince the doubters. Furthermore, the Knight and Kantowitz (1974) study suggested that the effect on the first response could be a significant fraction of the effect on the second response (perhaps 1/3 in one measurement). Countering this argument, it is possible that the effect observed by Knight and Kantowitz was due to response

grouping which is a phenomena that Pashler took pains to avoid. Thus, there doesn't appear to be a resolution without better experiments.

What is the central bottleneck theory of dual-task interference?

In the central bottleneck theory, one assumes there is a critical process that must be performed sequentially and not in parallel. This critical process creates a queue: multiple mental operations must line up in order to be carried out. Typically it is assumed that other processes before or after this bottleneck can be carried out in parallel.

A bottleneck in response selection or response production?

Pashler argues for the bottleneck occurring in response selection whereas others have argued for a bottleneck due to response production. Key to Pashler's argument are Case 3 and Case 4 of his predictions about factor interactions (p. 280-281). In Case 3, Pashler predicts that influencing the mechanism before a bottleneck will have no effect on the second response (R2), at least for short SOAs. This result was found when perceptual processing was manipulated by varying stimulus intensity (e.g. Pashler, 1984): No effect (or a subadditive effect) is seen in the response time to stimulus two. In Case 4, Pashler predicts that influencing processing at or after the bottleneck will have an effect on R2 which is independent of SOA. This result was found for factors that are assumed to manipulate response selection such as stimulus-response compatibility (McCann & Johnston, 1992). Pashler thus argues for a bottleneck which occurs at response selection.

Keele (1973) and others have suggested that the bottleneck occurs later in the processing associated with response production. An often cited result in support is from a study by Karlin and Kastenbaum (1968) that investigated the PRP with simple RT versus choice RT for the second response task. They found a subaddtivie effect of SOA and simple versus choice: There was a large difference between simple and choice with large SOAs but a small difference with small SOAs. Assuming that simple versus choice only affects response selection, this study suggests a bottleneck after response selection.

These two opposing results can be accommodated in different ways. The question is whether the failure of one or the other bottleneck theory was due to the misplaced bottleneck or due to faulty assumptions about the selective influence of the particular manipulation. In particular, Pashler suggests that the

Karlin and Kastenbaum result is due to the simple versus choice manipulation having much broader effects than simply extending the duration of the response selection stage.

On this point, Barry Kantowitz agreed that simple and choice RT are such different responses that it may be difficult to compare them in the way of Karlin and Kestenbaum. For example, temporal uncertainty has a different effect on choice versus simple RT. Barry suggested a 2 vs 4 choice RT experiment as a better experiment. According to Pashler, such an experiment was done by Hawkins, Church and de Lemos (1978) and did find additivity rather than subadditivity. In summary, this analysis of interactions was on the whole more consistent with a central bottleneck at response selection than at response production.

Might preparation be the main source of the PRP effect?

Tony Greenwald raised a question about the interpretation of preparation effects associated with the unpublished study of Ruthruff and Pashler described in Figure 6.3. They compared response times in experiments where subjects responded to a color task, a tone task, one or the other interleaved, or both responses. Pashler found slower responses to the both response condition compared to the interleaved condition. He argued that this difference could not be due to preparation. Tony didn't buy this interpretation. He didn't agree with the assumption that preparing to respond to one or the other of two stimuli is no different than preparing to respond to both stimuli.

Tony then described a PRP experiment in which two stimuli per trial are presented with a variable SOA. If a subject knows what the first stimulus will be, then the subject can prepare for the first stimulus initially and after the first stimulus appears, start preparing for the second stimulus. This would predict faster responses with increasing SOA due to the amount of time it takes to prepare. If the SOA is too short, the second task response may be slowed, or accuracy of the first task will suffer. If this is the case, subjects should show better performance on blocked SOA's with certainty of stimulus order. More generally, Tony argued that the PRP effect goes away for simple response time tasks if all temporal and event uncertainty is eliminated. This was presented as an example where subjects can program a single response sequence rather than two separate responses.

When does the PRP effect disappear?

The PRP effect is largely if not entirely eliminated in at least two situations. One case is when one of the tasks is to make a saccade towards a stimulus onset (Pashler, Carrier & Hoffman, 1993). The other case is when one of the tasks is to repeat a spoken word (shadowing, Greenwald & Shulman, 1973). In both cases, the other task was also a very compatible mapping between stimulus and response. Greenwald and Shulman suggested that the key factor is ideomotor compatibility which means that the stimulus is similar to the sensory feedback generated by the response. In the seminar, Tony Greenwald argued that these effects may be more consisted with a graded capacity account rather that a bottleneck account. Discussion then moved into the issue of the effect of practice and stimulus-response mapping, particularly with regard to playing music. Musicians in the group agreed that playing and performing some other task at the same time was very difficult unless the music was a well known, well practiced piece.

Hui-Lin Sarina Chien asked if any tasks didn't require response selection? For example, does shadowing not cause the PRP effect because it doesn't require response selection?

Beth Kerr raised the case in which PRP effect may be much reduced with a go-nogo response (Netick & Klapp, 1994). Again, the argument is that motor production or initiation may cause more interference than response selection.

The Meyer and Kieras alternative

We next took up Meyer and Kieras' (1992) suggestion of a different alternative to the response selection bottleneck. The authors suggest a flexible mechanism in which a particular task may result in a bottleneck whereas another task might avoid a bottleneck. In particular, they suggested that the typical PRP task instructions demand the first response be made as quickly as possible at possible cost to the second response. This is commonly done to avoid grouping of the two responses. Meyer and Kieras suggest that demanding an immediate response creates conditions that result in a bottleneck. They hint that other instructions should eliminate the bottleneck but perhaps not other kinds of dual-task interference. For the other side of this argument see Ruthruff, Pashler, and Klaassen (submitted). They pursued an alternative paradigm that encouraged response grouping. Even in this case, they found evidence consistent with a central bottleneck.

Choice of theory: Production system simulation vs mathematical model?

Another difference between the Meyer and Kieras work and previous work is their use of production system models rather than simple mathematical models of response time. Barry Kantowitz and John Palmer led a discussion contrasting simulations using production systems and analytic results based on mathematical models. The more complex simulations have the advantage of being able to describe very complex behaviors at the cost of the model not providing much insight into the process. Indeed, a successful simulation project often involves analytic models of its various special cases to both test the simulation and to help understand its workings. Thus, while complex simulations may be necessary for some problems, they need considerable elaboration to provide a satisfactory theory.

How can one apply ideas from the PRP effect to other situations?

Barry Kantowitz commented on the ease of overloading the human system with such a simple task as that of a PRP task. Barry finds that theories designed to account for the PRP paradigm are a good departure point for constructing a real world models.

Beth Kerr commented that the PRP paradigm leaves out too many important aspects of motor control. By not accounting for these factors, one cannot know the whole story about dual-task interference.

Survey of Attention
Week 7, Chapter 7: Questions on Attention and Memory
John Palmer
25 February 1999

Study Questions

1. What is the evidence for separate long-term and short-term memory systems?

2. What is the evidence for separate explicit (declarative) and implicit (procedural) memory systems?

3. What is the evidence against limited-capacity encoding in short-term memory?

4. What is the evidence for limited-capacity encoding in long-term memory?

5. What is the evidence for limited-capacity retrieval from long-term memory?

Discussion Questions

1. Pashler finds the evidence for distinguishing short and long-term memory systems more convincing than for distinguishing explicit and implicit memory systems. Do you agree? What future experiments might clarify these issues? Does this background information about memory help make the representational issues for memory explicit in comparison to the previous chapters on perception where the representational issues were largely implicit?

2. Pashler argues that short-term memory encoding and retention require no central capacity such as required for response selection. But in a previous section, he describes the sharply limited storage capacity of short-term memory. And others have argued (e.g. Shibuya & Bundesen, 1988; Palmer, 1990; Gegenfurtner & Sperling, 1993) that rate of encoding information into short-term memory has a limited capacity even before it reaches the storage capacity limit. How can these three different kinds of capacity coexist? Are they all related or are they three different aspects of short-term memory?

3. Consider the evidence for capacity limits in encoding into long-term memory. Are they compatible with a graded capacity or a bottleneck model (switching vs sharing)? How does this kind of capacity limit relate to others discussed above and elsewhere?

4. Review the experiments contrasting the capacity of explicit versus implicit memory. Is the case for similar capacity limits convincing? What further studies are needed?

The Counterpoint

Baddeley, Craik and others (e.g. Naveh-Benjamin, et al., 1998) have argued that the dual-task interference found for long-term memory encoding is larger than for retrieval. In contrast, Carrier and Pashler (1995) argue for a central capacity limit in retrieval that is related to response selection. How do the studies differ? Do the studies really conflict? What criticisms do each offer of the other? How can the disagreements be resolved?

Survey of Attention
Synopsis of Discussion
Week 7: Attention and Memory
Scott Tiernan and John Palmer
24 March 1999

Primary Reading

Pashler, H.E. (1998). The Psychology of Attention. MIT Press. Chapter 7.

Background Reading

Carrier, L. M. and Pashler, H. (1995). Attentional limits in memory retrieval. Journal of Experimental Psychology: Human Perception and Performance, 21, 1339-1346.

Counterpoint Reading

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. Journal of Experimental Psychology: Learning, Memory and Cognition, 24, 1091-1104

Initial Discussion

We started with a few minutes of discussion on a variety of topics. Mary Czerwinski remarked that she was surprised about Pashler's heavy use of the central bottleneck, which, she said he wrote about as if it were a known entity. Beth Kerr was surprised that Pashler never used the term "working memory." This lead to a brief discussion of the term "working memory", and Beth described it as "the director of the mind." Finally, Erin Harley initiated a discussion on the implicit/explicit distinction. Just how separate are they?

The experiments of Carrier and Pashler (1995)

Richard Abrams presented the Carrier and Pashler (1995) experiment on memory retrieval and response selection. The experiment tested the hypotheses of the central processing bottleneck versus parallel retrieval. If a central bottleneck holds for both memory retrieval and response selection, then

subjects will be unable to select a response while simultaneously retrieving a memory. On the other hand, if parallel processing holds, subjects should be able to retrieve and select at the same time. In this experiment, subjects memorized word pairs (e.g., bird-frog) and then did a series of PRP trials. In the PRP trials, subjects had two tasks: manual choice response to tones and to vocal cued recall of the paired associate. The auditory-manual task was always first and SOA between tasks was varied. Results supported the bottleneck hypothesis. The interpretation was that subjects were delayed in retrieving the word while they had to select a response to the first task.

Much of the discussion of the experiment revolved around the experimental method. It seemed clear that the effect was large. Mary Czerwinski remarked that an improvement to the experiment would have been to include a condition with reversed task orders. This prompted John Palmer to mention another of Pashler's papers in which he shows that two retrievals from two different categories has worst performance than two retrievals from the same category.

Tony Greenwald jumped in with a question about how much practice the subjects had in memorizing the word pairs. Richard wasn't sure, but mentioned that it must have been enough since error rates were low. Tony then moved on to the issue of making both tasks semantic. He suggested using paired opposites (e.g., black-white) such that the stimuli would be highly overlearned. John reminded us again of Pashler's study in which subjects showed better performance for dual retrieval when the items belonged to the same category rather than different categories. Tony felt that having subjects respond to "animal" would actually cause interference because the category "animal" contains so many items. Mary Czerwinski then brought us back to Tony's question about practice by asking the group whether or not learning reduces interference. Barry and John cited evidence from studies by Jim Johnston and others that training reduced the PRP effect but never made the effect go away.

Comparison to the Naveh-Benjamin et al. study

We then turned to the experiments by Craik's group which showed that in a concurrent dual task decreases performance in encoding but not retrieval. On the other hand, performance in the concurrent task was made worse by both encoding and retrieval. This prompted Barry Kantowitz to argue that there is a fundamental difference between memory and attention experiments. The Craik study is essentially a memory experiment while the Carrier and Pashler study is

essentially an attention experiment. Why would we expect the same outcomes? The dependent measures and many other details were different. Indeed, was it really the same kind of retrieval process in both cases?

Does rehearsal require capacity?

Continuing the contrast between attention versus memory, Barry Kantowitz turned to the topic of rehearsal. If you wanted to know whether or not rehearsal required attention, would you do an attention or a memory experiment? John prompted Barry for further explication on his conception of the difference between the two. "Would you do a Craik or a Pashler study to tell whether rehearsal is an attention demanding task?"

There were a variety of questions to clarify these issues. Beth Kerr commented that she thought Pashler saw rehearsal as intermittent and not requiring much attention. Sarina Chien asked more generally whether attention and memory were necessarily related to one another? Can we explain memory deficits in terms of attention? John's take on Pashler's thinking is that attention is general across domains. This does cause confusion for terms used in many ways in different contexts (e.g., capacity).

Barry Kantowitz jumped back in by stating that rehearsal is that ultimate memory manipulation. Speculating what would happen if attention were manipulated while doing rehearsal, he thought that we might see an effect on either memory or performance on an interpolated task. Mary Czerwinski felt sure that the answer was task-dependent, but that if it were rehearsed then she would expect to see no effect. Barry then responded by asking if that meant that rehearsal takes no attention?

As an example of a rehearsal study, Barry described a dual-task experiment involving short-term memory and a tapping task (Roediger, Knight, & Kantowitz, 1977). Subjects were required to maintain a modest memory load (e.g. memorize 5 words) and to tap a finger between two target locations specified by boxes. There was almost no effect of the tapping task on memory. Can we conclude from this that rehearsal requires no attention? Alternatively, rehearsal may require a different kind of capacity than required by the tapping task.

Beth Kerr suggested that Pashler would argue that initiating rehearsal and response selection does draw from a common central bottleneck, but that the

demands of reversal are so modest and intermittent that we don't observe their consequences. John tried to demonstrated that it was difficult to produce names of fruit while doing the tapping task. This prompted Tony Greenwald to suggest that the effect of rehearsal should show up in later retrievals of names of fruit since those required more capacity.

An applied example?

Barry Kantowitz then moved the discussion of rehearsal into the domain of applied research. Would you be worried, he asked, if your airplane pilot was busy rehearsing post-landing instructions during the landing process? After more background from Barry about what a pilot might be rehearsing while also attending to the landing, both Tony and Beth asked interesting questions. Tony asked: can one do unlimited rehearsal if it doesn't require capacity? People agreed that, even for rehearsal, we are limited by the short-term memory capacity. Beth asked about age effects and pilot performance. She remembered some literature showing that younger pilots performed better than older pilots, even with equivalent levels of experience. Barry replied with, "The world is set up to ameliorate these kind of memory effects."

Multiple kinds of capacity?

Tony pursued the theoretical discussion of rehearsal. If rehearsal doesn't affect other cognitive processes, does that mean that we posses different pools of resource (kinds of capacity)? Tony proposed a thought experiment in which one tries to remember 10 digits. In case 1, all 10 are presented visually. In case 2, five are visual and five are auditory. He proposed that since 10 may be more than 2 times 5 in terms of resource allocation, we should be better in the second case, when the 10 digits are processed with resource coming from two different pools. Barry reminded us of the effect of noise on decision making that Pashler mentioned early in the book: more combinations equals more noise, but not necessarily more capacity.

Structural versus strategic attention effects

Mary Czerwinski then asked where Wicken's multiple resource model fits in, which launched a discussion on structural versus strategic attention. Initially, a discussion of neural correlates of resource substitutions ensued. Tony Greenwald asked if it had been shown that a person can lose one type of capacity (e.g., auditory) without losing another (e.g., visual)? If this is the

case, then there must be a structural component. Erin Harley remembered a passage from the text (p. 329) that indicates that this is, in fact, the case. Barry Kantowitz didn't seem particularly convinced that "regions of the brain lighting up green" really proved much, since he feels that the brain is capable of doing things in so many different ways.

Attention and memory research methods

The group continued to ask about differences between memory and attention research in terms of method. For example, in perception and attention experiments one holds constant the sensory characteristics of the stimulus when manipulating attentional variables. What is the analog in memory research? Manipulating retention interval, reversal instructions, or the like while holding the stimulus to be memorized constant? Barry's thought was that in memory research one runs up against the problem of defining and manipulating stimuli, which is considerably easier in perception research. Barry also wondered what would happen if we did Pashler's experiment using ideomotor compatible stimuli?

Miscellany and summary

Sarina Chien asked whether or not memory was improved by trying harder? Most people felt that trying harder did not necessarily improve memory, but that changing strategies was often an effective way to improve memory.

Tony Greenwald commented that if we cannot defend multiple resource models, then in our research we should isolate the type of resource we are interested in, and study it exclusively. One difficulty, especially in applied research, is that you can't afford to run a new experiment for each new circumstance that comes along. We have to find something general that we can apply in multiple circumstances, even if it isn't right all the time.

We concluded with a general discussion of Pashler's book and this chapter in particular. John felt it laudable that Pashler tried to cover attention in a broader scope well beyond the common treatment that emphasizes visual attention. Mary Czerwinski and Tony Greenwald agreed, saying that they found the chapter helpful.

Survey of Attention
Week 8, Chapter 8: Questions on Automaticity, Effort, and Control
John Palmer
4 March 1999

Study Questions

1. What is the evidence regarding highly practiced processes not requiring capacity?

2. What is the evidence regarding highly practiced processes being performed without voluntary control?

3. What are capture errors (p 379)?

4. Under what conditions does increasing effort improve performance? When doesn't it?

5. Consider the experiments on task set. How do results of the cueing and blocking experiments compare? What are the implications?

Discussion Questions

1. Pashler argues again that laboratory tasks such as the PRP and visual search are better indicators of capacity demands than continuous tasks such as typing. Do his arguments suggest any testable predictions?

2. Pashler argues that despite the large effects of practice on PRP and complex search tasks, limited capacity remains. How does this idea coexist with the proposal that discriminating perceptual categories such as letters vs digits does not demand capacity? Similarly, shadowing does not demand the central bottleneck in the PRP paradigm. Must not one learn categories and language?

3. Pashler mentions in several places how practice can restructure a task. It can change the "response units" from single letters to words (p 363); it can change memory search from short-term to long-term memory retrieval (p 368); and, it can change an arithmetic algorithm of many steps to a single memory retrieval (p 376). With regard to the original unit in each example, this restructuring does reduce capacity demands. Perhaps automaticity is about using restructuring to avoid capacity limits?

The Counterpoint

The restructuring hypothesis has been raised in the past (e.g. Cheng, 1985) in a fashion similar to the shift from short-term to long-term memory retrieval described by Pashler (see Logan & Stadler, 1991). Schneider and Shiffrin (1985) responded in the counterpoint article that they agreed that such restructuring occurs. But they disagree that restructuring is the entire story. They interpret experiments on variable-mapped categories, transfer of training, and foil effects as evidence for consequences of practice beyond restructuring. To what extent do these arguments address Pashler's concerns?

Survey of Attention
Synopsis of Discussion
Week 8: Automaticity, Effort, and Control
Richard Abrams
10 December 1999

Primary Reading

Pashler, H.E. (1998). The Psychology of Attention. MIT Press. Chapter 8.

Background Reading

Logan, G. D. and Stadler, M. A. (1991). Mechanisms of performance improvement in consistent mapping memory search: Automaticity or strategy shift? Journal of Experimental Psychology: Learning, Memory and Cognition, 17, 478-496.

Counterpoint Reading

Schneider, W. and Shiffrin, R. (1985). Categorization (restructuring and automaticity: Two separable factors. Psychological Review, 92, 424-428.

Shiffrin, R. M. (1988). Attention. In R. A. Atkinson, R. J. Hernstein, G. Lindzey, and R. D. Luce (Eds.) Steven's Handbook of Experimental Psychology, Volume 2: Learning and Cognition. New York: Wiley (pp. 739-811).

Scott Tiernan presented Logan and Stadler's (1991) Experiment 1, which involved a memory search task. Subjects on each trial judged whether a probe item belonged to a set of items shown just before the trial. After a large number of trials, subjects received a single "catch trial" on which the probe belonged to the superset of items previously used as probes, but not to the set from that trial. The fact that many subjects produced false alarms on catch trials was interpreted as evidence against a process-learning model (the idea that improvement is due to progressively better comparison of the probe against an explicit representation of the current item set in short-term memory). If subjects had gotten better at making this comparison, then they should have been successful at rejecting the catch-trial probe. Not being able to reject the probe suggests that they had learned either to compare the probe

to the superset of probe items (superset hypothesis), or were responding on the basis of a learned category distinction between probes and distractors (category comparison hypothesis). We all agreed that the results strongly supported these latter two hypotheses, although there was probably not enough information to favor one over the other.

John asked whether we found these findings surprising: no one did. The consensus (to be elaborated later) that the kinds of hypotheses advocated by Logan and Stadler were more plausible than a parallel-processing interpretation as represented by the counterpoint article was succinctly expressed by Barry: "Comparing the probe to the superset is easier to believe than automaticity."

Barry wondered whether similar results would be obtained with visual search as opposed to memory search. John supposed that visual search would yield a somewhat different pattern of results, specifically that it would be likely to show more evidence of item-based learning and less of category-based learning (since categorization would probably be less easy in visual search). Barry then recalled a categorization study by Posner and Keele in which subjects practiced a categorization task, then did "catch trials" on which the items had not appeared before in the experiment but were possible exemplars of the learned categories. What would happen, Barry asked, if in the Logan and Stadler study the probes on the catch trials weren't actually items from the target superset, but exemplars (to a greater or lesser degree) of the target or distractor categories? More generally, to what extent does task improvement depend on stimulus-category mapping? Sue noted that when subjects are trained with letters as targets and digits as distractors, and then the category assignments are reversed, re-learning the new assignments becomes extremely difficult, suggesting that stimulus-category mapping is in fact critical in performance improvement.

Another of Barry's provocative questions: What does all this (improvement in set-size tasks) have to do with attention? John sought to answer by putting Logan and Stadler in its historical context. As we all know by now, the literature includes a great variety of set-size effects, which are rather difficult to interpret coherently. One approach to these effects is through purely attention-based models (Shiffrin is Pashler's prime example). Another approach is to focus on how practice diminishes set-size effects, without resort to concepts that involve attention. The Logan and Stadler study is relevant to attention exactly because it dispenses with attention, by showing that attention need not be invoked to explain phenomena typically interpreted in terms of

attention.

A discussion about restructuring ensued. With regard to Logan and Stadler's study, John framed the question of restructuring in terms of the information demands of the task. Classifying the probe is ostensibly an eight-bit task (given a superset of four possible targets and four distractors). But do subjects restructure it so that it becomes a one-bit task (target versus distractor categories)? Sue observed that this reduction in information appears to be not unlike chunking in memory. More generally, John argued that to understand the way in which performance improves on any task, we need to have a theory of that particular task. Several people then raised the question of whether this would be a theory based on capacity (and thus tied to the focus of discussion in earlier weeks) or a theory largely based on ideas of restructuring and strategy. Barry argued his belief that a theory based on restructuring would have similarities with the concept of strategic restructuring of workload (that is, the idea that people are adaptive and naturally rearrange elements of system architecture to better their performance). In fact, Barry suggested that the flexibility (even ingenuity) of people in restructuring tasks makes the study of capacity in complex tasks extremely challenging. Capacity can (possibly) be studied in the laboratory under controlled conditions using simple paradigms that have highly restricted outcomes. In more complex tasks that resemble real-world operations, however, capacity becomes elusive to measure because complexity affords subjects the opportunity to exercise their ingenuity in restructuring systems. What may look like an increase in capacity with practice in a complex task may actually be the result of restructuring and strategy. This difficulty in pinning down capacity in tasks other than restricted laboratory paradigms was considered by Barry to be a fundamental problem in psychology.

John took a slightly different point of view, namely that the reason subjects restructure tasks as they gain practice is to avoid capacity limitations. In this way, restructuring is closely related to capacity: to the extent that capacity imposes limits on performance, subjects will seek to improve performance through other means (restructuring). John highlighted this close link between capacity and restructuring by noting that Pashler's chapter could easily have been called "Attention and Learning." But we should not, added Barry, lose sight of the fact that restructuring can only go so far. There are fundamental capacity limitations that can't be overcome or circumvented by restructuring, as shown for example by the fact that nearly all PRP effects persist through even quite extensive practice.

Turning to the section of the chapter on effort, Erin asked how the concept of effort could be distinguished from the concept of attention. John reformulated the distinction as between effort and capacity, and Barry observed that Kahneman had developed a model that addressed the relation between effort and capacity explicitly. Kahneman found that by increasing effort, capacity can be increased, but at a diminishing rate. This led to the question of exactly how effort can be measured. Some experiments manipulate payoff and observe performance, assuming that effort increases with increased payoff. But effort can also be approached by assuming that some physiological index (pupil diameter, Fourier analysis of heart rate) represents the underlying construct of effort. Barry felt that Kahneman had used this latter approach effectively, that Kahneman was on the right track with this approach and that it didn't catch on only because it conflicted with the Zeitgeist. A physiological approach might be more viable now, because of the development of interest in localizing specific cognitive operations to brain areas (an interest that explains the current popularity of imaging).

Effort can also be measured subjectively. Subjective measures of effort are compelling, because everyone is familiar with what effort is and has had experience with different kinds of effort, and describing effort in terms of magnitude or scale is something people do in everyday discourse. Barry cautioned, however, that this familiarity doesn't necessarily translate into accuracy. For one thing, measuring effort subjectively is susceptible to the same kinds of problems that have made introspection (of other psychological states) a generally discredited method. And empirically, the unreliability of subjective measures is shown at an extreme by the fact that there are often dissociations when effort is measured both objectively and subjectively. In general, subjective workload measures are practical when the underlying processes are uniform and simple, but problems arise when tasks are complex and presumably more than a single, simple kind of effort is involved.

Survey of Attention
Week 9, Questions for the Synopsis
John Palmer
9 March 1999

Perception Questions

1. Might the ease of selection be determined solely by the discriminability of the attribute rather than if it is "physical" or "semantic"? (Palmer)

2. Pashler presents a fairly convincing case for controlled parallel processing as an alternative to the early and late selection theories. Are we justified in "reject[ing] both theories" as he writes in the conclusion? (LeeTiernan)

3. How good is the evidence that there are no capacity limits for simple perceptual categories (e.g. letter vs digit)? Might the cases with no observed capacity limits be due to using simple features to discriminate the categories? If so, does this make early selection a viable theory for attention in perception? (Recall that the lack of capacity limits for simple perceptual categories was critical to rejecting early selection theory.) (Palmer)

Motor and Memory Questions

4. Pashler (Chapter 6; page 286) acknowledges the possibility of a second bottleneck linked with initiating motor commands (he considers this bottleneck an effector-dependent bottleneck in response execution). How important is this limitation? Does it overshadow any potential bottleneck in response-selection when it comes to performing most applied (real world) tasks? Is this limitation really effector dependent? Might it be a manifestation of structural defaults that favor temporal and spatial coupling in responses? (Kerr)

5. How can one distinguish how much of the PRP effect is due to preparation or to a central bottleneck? (Palmer)

Automaticity, Effort, and Control Questions

6. What behaviors do you get for free? Should we think about going after "getting something for free" in design of tools? How far can that take that really take us? (Risden)

7. How can we explain the phenomenon of applying greater effort in a task and ending up performing worse than had we applied less effort? This question relates to automaticity, effort and the bottleneck proposal. At the beginning of the book (chap 1), Pashler describes that for certain highly automatic tasks such as tieing shoes, adding additional effort reduces one's performance at such tasks. Why should we do worse when we apply MORE effort for this special case of "superskills"? Could attention to or effort for tieing your shoe be a source of interference? How can this phenomenon help to better define the bottleneck argument or our ideas about capacity? Have there been any studies on this subject which we could discuss? (Hartzler)

General Selection Questions

8) How much individual strategy is involved in attention. For example, can I choose to use switching techniques to do two tasks simultaneously? Alternatively, can I choose to attend only to one stimulus and effectively prevent processing of other stimuli, or is parallel processing to some degree "automatic" and out of my control? (Harley)

General Capacity Questions

9. What is the definition of "capacity", and what are the varieties of capacity that are required to summarize what we know about attention? (Czerwinski)

10. How do the different structural effects (e.g., interference or lack thereof for different kinds of information--visual/spatial, verbal, motor, etc.) influence our definition of capacity? (Czerwinski)

11) What should I take away from this book about the issue of serial versus parallel processing? (Harley)

12. The term capacity is used to describe several different mechanisms of processing dependence. These include the central bottleneck, serial vs parallel processing more generally, and dependencies among parallel processes. Would it be useful to limit the use of the general term "capacity" and focus on these more specific hypotheses? (Palmer)

Even More General Questions

13. In many cases in this book, methodologies have been open to alternative explanations to the extent that they appear to be "boxed in" (i.e., it seems like it will be hard to make progress in any direction). Do you think this field is ready for a new direction that will stir things up and if so what do you think that will be? (Risden)

14. What does this book/this body of work tell us about specific directions for research on how technology can better leverage human capabilities to make user interfaces better tools (e.g. easier to use, extends our cognitive abilities, etc.)? (Risden)

15. Are there any general strategies on how to determine if a given phenomena is due to attentional mechanisms rather than other processes such as sensory interaction, memory association, learning, etc.? (Palmer)

16. Many arguments in attention depend on assumptions about the structure of the relevant processes as well as those about attention. For example, the analysis of memory assumed the distinction between short-term and long-term memory; the analysis of the PRP effect assumed a sequence of separate processing stages and certain factors that separately influence each stage; and, the analysis of perceptual capacity assumed that the stimuli corresponded to the relevant perceptual representations. The question is what aspects of the processing and representation must be made explicit in order to evaluate the issues of attention? (Palmer)

Survey of Attention
Synopsis of Discussion
Week 9: Synopsis
Erin Harley and John Palmer
10 December 1999

Overview

The bulk of this session was going through the proposed questions for Hal Pashler visit the following week. This part of the discussion was not recorded. Below are other comments that arose during the discussion.

PRP

How can we further investigate the issues of the PRP effect? Pashler and Ruthruff forced subjects to group their responses and then looked at the delays in grouping as a function of the SOA. Another good technique is to use an SOA of zero. What are other ways of manipulating the response? If you map all the stimuli to one response you lose the effect. There is still stimulus uncertainty but now there is no longer any response uncertainty.

Is Pashler's book biased?

All agreed that the book has biases, so the question became just how biased is the book? Barry described two types of biases. The first, less serious type, is when an author does not consider all of the relevant and salient data. This can sometimes be a useful bias. The second, more serious bias, is when an author chooses to include discussion of past theories but then does not present all past theories that show merit. Barry felt that Pashler gave insufficient consideration to Broadbent's latest model, the multiple resource model, and Kahneman's model.

The book is selective. If you are confused about the theoretical building blocks of attention you will be confused about Pashler's conclusions. Some people in the group felt that he could have made the same points in a different, and better, way. While Pashler tries to reject theories, Kahneman tried to make the case for his theory. Could this be a result of Pashler trying to choose the most convenient method rather than the best method? Would you rather be precisely wrong or approximately right?

Capacity

Should we limit our use of the term capacity? How should we define capacity? Channel capacity (telephones, switchboard) was a well-defined and clearly articulated concept, but capacity cannot exist at this autonomous level. Perhaps Broadbent had the right idea in using the term capacity in the broadest sense.

There is an expectation in the field of psychology that at least some tiny portion of the research we do will make a difference in the real world. Does Pashler's controlled-parallel model determine how you would design an air traffic control tower? Simpler models could make the same predictions and be far more useful in the real world. Are there really true differences in the predictions made by these models, or are we arguing over the 5th decimal place. Would Pashler's book increase funding to the field of Psychology?

Coupling

We know that manual and vocal responses can easily be grouped, but our system puts temporal constraints on the degree to which we can do two manual tasks simultaneously. Split-brain patients (patients whose corpus callosum has been cut) can draw a circle with one hand and a square with the other hand and show no signs of interference. The default of the system seems to be to have spatial and temporal coupling; it appears as if the system needs this coupling. Pashler's book is lacking in his discussion of motor control.

Is there a psychology of attention?

Does attention exist as its own entity or do we have attention subsumed under each modality? Can visual attention tell us anything about auditory attention? If there is not one central system controlling attention then we cannot make generalizations about attentional processes across modalities. To answer these questions we must look for evidence of independence or dependence between attentional processes in different modalities. Pashler found evidence for dependence between memory and motor. It is also possible that there is a higher order emergent level which controls the whole thing. Modularity would exist but it would have emergent levels. Characteristics of an emergent system would not be guessed by looking at the simple components.

Final thoughts

To what extent is adaptation voluntary? When does attention become involuntary? How do people restructure in long term memory? Is the PRP effect a narrow paradigm or is it a true representation of human ability?

Survey of Attention
Week 10, Questions for Pashler
John Palmer
29 March 1999

Specific Questions

1. Might the ease of selection be determined solely by the discriminability of the stimulus alternatives rather than if the relevant attribute is "physical" or "semantic"?

2. How good is the evidence that there are no capacity limits for processing simple perceptual categories (e.g. letter vs digit)? Might the cases with no observed capacity limits be due to using simple features to discriminate the categories? If so, does this make early selection a viable theory for attention in perception? (Recall that the lack of capacity limits for simple perceptual categories was critical to rejecting early selection theory.)

3. Pashler (Chapter 6; page 286) acknowledges the possibility of a second bottleneck linked with initiating motor commands (he considers this bottleneck an effector-dependent bottleneck in response execution). How important is this limitation? Does it overshadow any potential bottleneck in response-selection when it comes to performing most applied (real world) tasks? Is this limitation really effector dependent? Might it be a manifestation of structural defaults that favor temporal and spatial coupling in responses?

4. How can one distinguish how much of the PRP effect is due to preparation rather than a central bottleneck?

5. The work on attention and memory appears to be in a state of flux. Could you summarize recent developments in your own work and that of other researchers such as Craik and Jolicoeur?

General Questions

6. What behaviors or component processes do you get for free (with unlimited capacity)? Should one think about "getting something for free" in the design of software and other tools?

7. What is the definition of "capacity"? The term capacity is used to describe several different mechanisms of processing dependence. These include the central bottleneck, serial vs parallel processing more generally, and dependencies among parallel processes. Would it be useful to limit the use of the general term "capacity" and focus on these more specific hypotheses? For example, Townsend (1974) uses the term capacity only to refer to dependencies between component processes and not to the issue of serial vs parallel processing.

8. It is clear how the viewpoint of this book contrasts with Broadbent (1958) because of the detailed discussion of his version of early selection theory. Less clear is the comparison to Broadbent (1971). What are the similarities and differences between the viewpoint of this book and Broadbent (1971)?

9. This book makes an implicit case that the phenomena of attention are similar in vision, audition, memory, and motor control. This view is in contrast to one in which the processes of each domain is more modular. How can one make the arguments against modularity more explicit?

10. What does this body of work tell us about profitable directions for applied research in human factors? For example, what research might inform one of how to make better user interfaces (e.g. easier to use, extend our cognitive abilities)?

Survey of Attention
Synopsis of Discussion
Week 10: Pashler's Presentation
Erin Harley and John Palmer
19 April 1999

Overview

Hal Pashler came up to Seattle to discuss his book at the last meeting of the seminar. The outline of the discussion followed the list of questions we had formulated at the previous meeting. Other issues are interspersed in this description in the order they came up.

Selection and Discriminability

Question 1: Might the ease of selection be determined solely by the discriminability of the stimulus alternatives rather than if the relevant attribute is "physical" or "semantic"?

In other words, could discriminability be the whole story? Location and color are easy to select, but they are also highly discriminable. It would be useful to clarify whether or not one is talking about physically discriminable or semantically discriminable. "Tiger" and "Lion" are semantically discriminable but it would probably be impossible to come up with two words that are as highly discriminable as the colors red and green.

Tony Greenwald asked how are discriminability and selection to be distinguished. Palmer used the study of Smallman and Boynton (1990) as an example of measuring the ease of selection. They asked subjects to locate a particular shape among all of the red stimuli in a display. When the red was highly discriminable from other colors in the display, performance was no different between conditions that contained other colored shapes and conditions that contained only the red shapes. Thus highly discriminable colors allowed essentially perfect selection. Less discriminable colors allowed less perfect selection. This is a distinct measurement from simple discriminability that can be measured with a single stimulus.

Pashler suggested that if there is only a single possible target in a search task, then physical discriminability is the key. However, if the target can be any item from a set or a category, then semantics plays a large role in the ease of specifying the category to be selected. He provided an example of this using letters and digits. In Task A, subjects are asked to search for a number between 1 and 5 among a distractor set of 26 letters. In Task B, subjects are asked to search for one of 5 target characters (some letters and some numbers) among 26 distracters (some letters and some numbers). In task A the memory load is held constant: all the targets are from a well-learned category while all the distracters are from a different well-learned category.

One is able to search for one member of a set without much more capacity than searching for one item if the set is well practiced or well learned. In task B, the lack of a semantic structure has a large effect because one must looking for a complicated set of items.

Tony Greenwald described an experiment of his in which subjects are asked to pick out flower names from a list of words. In one condition these words are pleasant words and in another condition they are unpleasant words. Subjects have greater difficulty picking out the flower names from a list of pleasant words. This is likely due to the fact that there is a large degree of overlap between the semantic categories of flowers and pleasant words making the task difficult.

In closing, it is clear that semantics will play a role in selection under some circumstances. However, the narrower question intended here wasn't made clear. Would the ability to select by a semantic category such as letter vs digit in a search task be predictable from the discriminability between single letters and digits? Or, is there an additional decrement in performance due to the processing of semantic distinctions being "later" in the processing stream than physical distinctions? It seems that any early selection theory would predict the later. While a late selection theory would predict that discriminating the category would be similar to the worst of the exemplars. However, this letter versus digit comparison has alternative interpretations that are discussed below. Further experiments examining discriminability more directly haven't been done because they require one to match discriminability across different conditions.

Capacity limits for Simple Perceptual Categories

Question 2: How good is the evidence that there are no capacity limits for processing simple perceptual categories (e.g. letter vs. digit)? Might the cases with no observed capacity limits be due to using simple features to discriminate the categories (e.g. Duncan, 1983; Krueger, 1984)? If so, does this make early selection a viable theory for attention in perception? (Recall that the lack of capacity limits for simple perceptual categories was critical to rejecting early selection theory.)

Pashler thought the best argument was based on an experiment in which subjects were asked to report the highest number in a display of digits (Pashler and Badgio, 1987, Attention and Performance). The idea is that this task requires the perception of the meaning of the number in a way that typical search tasks do not. This experiment used the simultaneous/successive paradigm developed by Eriksen and Spencer (1969) and Shiffrin and Gardner (1972). For relatively short SOAs between successive displays (<100 ms), Pashler and Badgio found no advantage for successive displays in the higher digit task. Thus, they argue that there is no rapid serial processing to circumvent limited capacity of a simultaneous display.

Pashler also described unpublished followup experiments to distinguish whether people

used semantic cues to do the task or selected the number based on some simple physical features. For example one could first search for a nine, then an eight, etc. until the highest number was detected. (add description of results ???) Results of this experiment indicate a parallel semantic analysis for numbers: Subjects were very fast at detecting the highest number and it did not appear as if they were cycling through physical features of each number one at a time.

One concern that comes to mind after the discussion is whether these results would hold up for long SOAs such as the 500 ms used by Shiffrin and Gardner. Perhaps there is never rapid serial processing on the order of 50 ms, but still limited capacity for identification. This would show up for longer SOAs that allowed the slow serial process to take advantage of the sequential display.

Palmer asked if there is any other evidence against feature search? Pashler suggested the next best would be experiments with the mixing of font types (xxx).

Motor Production Bottlenecks

Question 3: Pashler (Chapter 6: page 286) acknowledges the possibility of a second bottleneck linked with initiating motor commands. He considers this bottleneck an effector-dependent bottleneck in response execution. How important is this limitation? Does it overshadow any potential bottleneck in response-selection when it comes to performing most applied (real world) tasks? Is this limitation really effector dependent? Might it be a manifestation of structural defaults that favor temporal and spatial coupling in responses?

Pashler described a study of this response execution bottleneck with a PRP experiment in which Task 1 required one or three key presses and Task 2 required a vocal response (Pashler and Christian, 1996, unpublished?). Task 2 comes while subjects were in the middle of Task 1 and subjects showed no difficulty performing the vocal response in the middle of the key presses. Having three key presses did slow the onset of the vocal response by about 20 or 30 ms more than having only one key press, but this is a very small amount compared to the effect expected for a response bottleneck. When Task 2 was switched to a manual response instead of a vocal response, subjects were no longer able to give response 2 during response 1. They had to wait until the completion of response 1 before initiating response 2. Thus, the effect was much larger (in the 100s of ms?). This experiment was also tried using a foot response for task 2 and the interference was still present, that is, the foot response could not be made until the manual response was complete.

Does this manual-manual interference reflect a production bottleneck or a response bottleneck? Two manual responses may require one to finish response 1 before initiating the selection of response 2, but Pashler thought that was unlikely. It appears that subjects have already chosen response 2 before they complete response 1 because as soon as response 1 has

been produced, response 2 is made with no delay. The second response is just waiting to have access to manual production. All of this overlap indicates two separate bottlenecks, a response bottleneck and a production bottleneck.

Beth Kerr argued that we can almost never use our hands independently. We also link vocal with manual responses easily such as in eating. Perhaps the exception found between manual and vocal tasks has to do with physical and linguistic materials? Would you also avoid the response production bottleneck if the vocal task was not using linguistic materials. For example making noises of selected pitch and loudness?

How about when manual responses involve linguistic materials? Signers use their two hands differently to sign; however, they are signing the same conversation, sentence or word. Could they simultaneously sign two different conversations? Can percussionists simultaneously play two different rhythms? Although percussionist are able to play different sub-rhythms with their two hands, it is likely that in order to do this they must internally choose one common higher rhythm to combine the two sub-rhythms. Another question is whether one could do a non-linguistic task with your feet while you signed with your hands?

What do these arguments tell us about driving and talking on a cellular phone simultaneously? Perhaps the physical demands of dialing and holding the phone would greatly interfere with driving, but speaking on the phone would not.

Response Selection and Preparation

Question 4: How can one distinguish how much of the PRP effect is due to preparation rather than a central bottleneck?

In Pashler's book, he evaluates the contribution of preparation in an unpublished study with Ruthruff (p 275-576). They compared performance with single tasks, two-interlaced tasks, and dual tasks. The argument was that one had to prepare both tasks in the interlaced task condition but only had to execute both tasks in the dual task condition. Thus, the comparison between single vs interlaced conditions reveals preparation effects and the comparison between interlaced vs dual task conditions reveals dual-task interference above and beyond preparation effects. Both of these comparisons revealed effects on the order of 100 ms. Thus, they argue that both effects play a role, but preparation isn't sufficient to account for all of dual-task interference.

Tony Greenwald objected to comparing a dual task with two interlaced tasks. He argued that preparation demands are different for any single task and a dual task: the subject knows he/she will have to prepare two responses rather than one. Is Pashler underestimating the contribution of preparation? Pashler didn't think so. He argued that when you are given only one task when you are expecting two, you think to yourself, "where is the other one,". Thus, you are prepared for both tasks.

What is an example of preparation? Rehearsing the mapping of a response can be thought of as preparation. Tony explained that in his own experiments when he wants to go fast he has to think ahead about what the stimuli will be and what the appropriate responses will be for each of those stimuli.

Maybe one could prevent preparation by asking subjects to keep their minds clear before the onset of the stimuli. Another way to bypass preparation issues might be to use serial RT experiments where there is only one mapping but many stimuli (e.g. Pashler, 1994).

Attention and Memory

Question 5: The work on attention and memory appears to be in a state of flux. Could you summarize recent developments in your own work and that of other researchers such as Craik and Jolicoeur?

Perhaps the word "flux" should be replaced with "disarray." By the analysis of Carrier and Pashler (1995), memory retrieval is the most capacity demanding kind of process. You can only do one memory retrieval at a time. In contrast, Craik claims that memory storage is capacity demanding while retrieval is not. To support this view, Craik did a study of word recall and found little interference in word recall when subjects were required to perform a second task simultaneously. The problem with this study, as Pashler explained, is that there was too much slack time during the second task allowing subjects to do a lot of switching. Pashler theorizes that experimenters who find little interference between multiple memory retrievals are underestimating how much time is stolen in between tasks. Craik found that the concurrent task suffers from retrieval but that the retrieval does not suffer. This is consistent with subjects stealing time. They are testing this hypothesis by repeating Craik's measurements with a preview task that minimizes the slack time in the concurrent task. One might also try and test this hypothesis using a tracking task as the concurrent task and see if subjects get jerky.

It was commonly believed that it is possible to put things into short-term memory (STM) without any interference, but it is not possible to get things in or out of long-term memory (LTM) without interference. This view was reinforced by the study of Pashler (1993) described in the text. Now, however, there is evidence that interference does occur for STM under at least some conditions (e.g. Jolicoeur and Dell'Acqua, 1998). Pashler presented the following view to accomadate both results. If you think of STM as an executive with many slaves, then you can imagine that the slaves can hold information while the executive is actively making decisions (no interference). Alternatively, if all of the slaves are being used the executive can also be used to hold things but then it is unable to make decisions. When the latter case occurs, you will find interference when trying to put things into STM. Thus, when the executive is busy, STM processing may require the central bottleneck. Otherwise, the executive can act independently.

In Pashler's book, there was an experiment described on page 347 in which subjects had to retrieve items from some category (Rohrer, Pashler, and Etchegaray, 1998?). Tony Greenwald pointed out that in such a task not only does one have to generate new retrievals but one must also remember which items have already recalled to prevent from recalling the same items a second or third time. Because of this additional demand, the task becomes increasingly difficult as more words are recalled. How could one eliminate this additional effect on memory? Perhaps subjects could count backwards, recalling numbers but not needing to remember which numbers had already been recalled aside from the most recently recalled number. Another possibility would be to have subjects recall words beginning with the last letter of the previous word, or one could cycle through various categories and ask subjects to give one exemplar of each category.

Does reading require retrieval from LTM? Pashler answered that perhaps shallow comprehension does not, but deeper meaning does. What about single words, even nonsense words? Do reading these require retrieval from LTM? Here the evidence of minimal interference between perceptual tasks and response selection tasks suggest that they don't require retrieval in the same way as explicit retrieval tasks.

Memory Systems

The group asked Pashler to explain his opinions about the distinction between implicit and explicit memory. Pashler is not convinced that implicit and explicit memory are two separate memory structures. He suggests that perhaps the difference between the two is simply that implicit memory requires less retrieval than explicit memory. He argued that in this case there is no clear-cut evidence for distinct memory systems. Simply because activation may occur in different areas of the brain does not necessarily indicate different systems at work. A library has multiple floors, but that does not mean that the books on one floor are on a different system than the books on another floor. For this reason one should be cautious about using brain localization arguments to identify information processing systems. In these arguments, if two tasks produce activation in two different areas of the brain, they are said to be separate systems. This is not necessarily the case.

Pashler argued that the STM-LTM distinction is better established than the explicit-implicit distinction. A variety of further questions were raised: If iconic memory is continued cell firing, could STM simply be even more continued cell firing? Perhaps repetition priming is just a manifestation of LTM, a kind of association of LTM. However this does not explain priming of a semantic associate. Subjects show more rapid forgetting for recall and recognition memory than for priming and implicit memory, but maybe that is the case because there is less interference for priming and implicit memory. If one learns to do a mirror-tracing task, what kinds of information will interfere with that knowledge?

Relevance to Applied Research

Question 6: What behaviors or component processes do you get for free (with unlimited capacity)? Should one think about "getting something for free" in the design of software and other tools? (This is also relevant to Question 10 that was not directly discussed.)

To illustrate the question, Palmer argued that the Graphical User Interface (GUI) is a response by programmers to limit the number of memory retrievals a user must make when using a computer program. People in user interface and software development talk about leveraging pre-attentive capabilities. What are pre-attentive capabilities and how can one use our knowledge of attention to improve computer software and GUI's? Color and location are probably the most fundamental pre-attentive features; one "get them for free," but what balance of pre-attentive and semantic features should be used in GUI's. Will people be drawn to relevant objects or just to pre-attentive features?

These issues depend on the details of the task. For example, if all the buttons are colored differently they will be easier to discriminate, but one may lose the effect of color highlighting for items in the display due to the increased clutter of many colors.

It would not be a good idea to limit GUI's solely to pre-attentive features because doing so increases memory load demands. For example, words are not pre-attentive but if buttons on a GUI are not labeled then the user has to remember which colors, shapes and locations correspond to which things. This would become very demanding for the user.

Capacity

Question 7: What is the definition of "capacity"? The term capacity is used to describe several different mechanisms of processing dependence. These include the central bottleneck, serial vs. parallel processing more generally, and dependencies among parallel processes. Would it be useful to limit the use of the general term "capacity" and focus on these more specific hypotheses? For example, Townsend (1974) uses the term capacity only to refer to dependencies between component processes and not to the issue of serial vs. parallel processing.

Palmer elaborated the question by comparing the three ways that capacity is used in the literature (Palmer, 1998). Broadbent uses it to refer to the entire system from input to output. Pashler and many others refer to the capacity of a particular stage of processing such as perception, response selection, or memory retrieval. Finally, Townsend uses capacity to refer to single component processes that may be arranged in either serial or parallel configurations. Thus, Townsend used the concept of serial vs parallel

independently of the concept of capacity. This is different from either Broadbent or Pashler's usage in which a serial process would typically imply limited capacity because the serial process limits the information conveyed per unit time.

Pashler asked for a clarification on the concept of serial processing without capacity limitations. Palmer illustrated three types of serial processing with the following figure that indicates duration by the length of the horizontal "box".

Time -->

A) |____| single process of given duration,

B) |____| |____| two serial processes, unlimited capacity,

C) |_____| |_____| two serial processes, limited capacity,

D) |__| |__| two serial processes, super-capacity.

This figure illustrates Townsend's use of the term capacity to refer to the dependence of a single elemental process on the presence of other processes in the set. Case A is the baseline of a single process of some given duration. Case B is when two processes are sequential, but the duration is the same as when they occur individually. Thus, the duration is independent of the number of processes and is referred to as unlimited capacity. Case C is when the two processes are sequential and take longer than when they occur individually. Thus, the duration depends on the number of processes and is said to have limited capacity. Finally, Case D, is when the two processes are sequential and have a shorter duration than when they occur individually. This is an example of super capacity.

The discussion then moved onto examples of these kinds of combinations of processing. Certainly serial and unlimited capacity (in Townsend's usage) is the case in most common examples. But one can find examples of serial and limited capacity. Pashler came up with the case we discussed earlier: a response selection bottleneck that implies serial processing with the additional constraint of being able to prepare only one process at a time. This would result in a serial process with limited capacity. Note, however, that preparation effects are not what first comes to one's mind when discussing capacity. This example helps make clear how the term capacity has been generalized from Broadbent's initial idea.

After this discussion, it is still not clear to me (Palmer) that using the term "capacity" is a good idea. In particular, capacity is being used in a very general sense but is burdened with specific connotations from its origins in information theory. Perhaps it would be better for the general case to use terms such as "mulitple-task interference" or "multiple-stimulus interference". Then one could consider alternative hypothesis for the source of this interference

effect. These would include the original Broadbent idea of graded limited-capacity channel, an all-or-none serial processing bottleneck, and many other possibilities such as the preparation effect discussed above. The use of a term such as interference would make it easier to discuss non-attentional explanations of these effects such as lateral masking in the visual search situation.

An Interpretation of Serial Processing

How can one think about these serial-processing models? Perhaps the brain is built in such a way that the pattern of activation that allows people to comprehend input into one ear precludes activation of input into the other ear. Activation in a structure can be incompatible with another pattern of activation in that same structure. If this is the case, then one should expect to see evidence of a serial processing.

Pashler (ref?) described an experiment in which subjects were asked to press one key if they saw an 'A' or heard a high tone, and press another key if they saw a 'B' or heard a low tone. People were faster when they saw the letter and heard the corresponding tone together than when they just saw one letter or heard one tone alone. The idea is that if two look-ups leads one to the same place then one can do them simultaneously. This makes sense if you think of these look-ups as activation patterns. Two cues can move you into one state faster than one cue alone, but they cannot move you into two states, nor can one cue move you into two different states. With enough practice people can couple two responses to a single cue, but once these responses have been coupled then you are essentially only doing one look-up.

This lead to a discussion of alternative word meanings in sentence context. Both meanings of an ambiguous word get activated early on in processing but then one gets inhibited. Are these multiple retrievals and subsequent inhibition demanding? Good comprehenders have been found to be good at inhibiting while poor comprehenders are not. The sentence "I am sitting on the riverbank" primes both "money" and "water," but if one allows a certain amount of time to pass then only the correct meaning is primed (in this case "water"). Tipper argued that selection, even by space, is due to inhibition. Pashler did not agree with this view but he does agree that there is a great deal of evidence for such effects in comprehension.

Historical Background

Question 8. It is clear how the viewpoint of this book contrasts with Broadbent (1958) because of the detailed discussion of his version of early selection theory. Less clear is the comparison to Broadbent (1971). What are the similarities and differences between the viewpoint of this book and Broadbent (1971)?

While this question was not discussed in the seminar, Pashler and I

(Palmer) discussed it informally. In his book, Pashler did not intend to make any historical comparisons between his book and other recent books on attention. Indeed, he has not thought recently about Broadbent's 1971 book. We will have to make these comparisons for ourselves.

Modularity of Attention

Question 9. This book makes an implicit case that the phenomena of attention are similar in vision, audition, memory and motor control. This view is in contrast to one in which the processes of each domain are more modular. How can one make the arguments against modularity more explicit?

There were several observations about these issues. First, can one ask the same questions in all of these modalities? It is quite difficult to do just that. For this point of view, the book does succeed in unifying the topics of attention by posing the same questions across a variety of domains.

Second, Allport and others have suggested that specific aspects of processing might have different attentional regularities and mechanisms. Thus, attentional phenomena for perception may have little to do with the attentional phenomena of memory or motor control. Here the results are more complex. There is little evidence of common mechanisms between perception and motor control. But, there is evidence for common mechanisms between motor control and memory. Thus there may be some modularity, but not specifically of the sort suggested on the basis of prior theories.

A third observation was generated from a study of Pashler, Luck, Hillyard, Mangun, O'Brien and Gazzaniga, 1994). Split-brain patients were given two tasks each mediated by a separate hemisphere. For each task, the sensory input and motor output where directed to one hemisphere. Under these conditions, there is still a PRP effect much like with normals. Such interference must be mediated by subcortical structures. Thus, here is a case in which the "modularity" of the split brain patients is refuted for the case of dual-task interference.

Survey of Attention
Summary of generalizations
John Palmer
29 March 1999

Selection and the Unattended in Perception

1. No direct memory for unattended stimuli.

2. Selection depends on physical discriminability.

3. Selection affects indirect memory for unattended stimuli.

11. More stimuli reduce this indirect memory for unattended stimuli.

13. Possible to select visual and auditory information from different locations.

A15. Selection can be influenced by particular stimuli as well as by voluntary intentions.

A16. Switching attention may take at least 200 ms.

Divided Attention in Perception

4. Unlimited capacity for processing features.

5. Unlimited capacity for processing simple categories (e.g. letter vs digit).

6. Limited capacity for processing complex categories (e.g. words).

7. Limited capacity for processing multiple targets (two-target effect).

12. Capacity limits apply to attended stimuli, not all stimuli.

14. Capacity limits are more severe within a modality compared to multiple modalities.

Attentional Set in Perception

8. Orthogonal cuing benefits in noise only.

9. Identity cuing usually affects bias and not sensitivity.

10. Identity cueing does affect sensitivity for long duration displays and when cues indicate large changes in the necessary decisions.

Attention in Sensorimotor Tasks

A17. Strong interference between two unrelated speeded responses close in time (PRP interference).

A18. The PRP interference can be much larger for the response to the second task compared to the first task.

A19. An analysis of the factor interactions generally supports a bottleneck model related to response selection rather than perception or response production.

A20. Similar response modes cause larger PRP interference than dissimilar modes.

A21. PRP interference is increased by increasing response uncertainty (e.g. S-R mapping).

A22. Perceptual tasks with a delayed response do not produce PRP interference with a single speeded response.

Attention in Memory

A23. With the PRP paradigm, long-term memory encoding and retrieval interfere with making speeded responses.

A24. Such PRP interference may not occur with short-term memory encoding and retrieval.

A25. Short-term memory has other kinds of capacity limits in encoding and storage.

Attention in Learning

A26. Practice can have large effects on absolute performance and on some measures of capacity.

A27. In experiments that minimize the role of restructuring, there is little evidence of practice modifying measures of capacity.

A28. Demonstrations of practice affecting the degree of voluntary control can be accounted for by either restructuring or automaticity.

Theoretical Generalizations

A29. Reject early selection models (e.g. Broadbent, 1958).

A30. Reject late selection models (e.g. Duncan, 1980).

A31. Support the controlled parallel processing model (Pashler, 1998).

A32. Reject a graded capacity account of central limitations (e.g. Kahneman, 1973)

A33. Support a central bottleneck model of central limitations (Pashler, 1998).

A34. Perceptual capacity and the central bottleneck are distinct.

A35. The central bottleneck may be necessary for long-term memory encoding and retrieval but not short-term memory.

A36. The effects of practice are often due to restructuring that replaces processes with limited capacity with processes that avoid these limits.

A37. There is little evidence that practice results in automatic processes rather than restructuring. By automatic, we mean processes that use less capacity and are performed with less voluntary control.

Note: Numbers and numbered statements from Pashler; organization, text, and "A" statements from Palmer.

Survey of Attention
Questions for Future Research
John Palmer
29 March 1999

From Pashler (more or less)

1. What is the nature of perceptual capacity limits for processing complex categories?

2. What is the nature of the representation upon which selection is based: multiple objects versus sets of attributes?

3. What is the nature of the central bottleneck?

From Palmer

4. What is the quantitative nature of selection (e.g. listening bands) and how does it relate to physiology?

5. Do different paradigms demonstrate that the same complex perceptual tasks require limited capacity (e.g. search, dual task, sim/succ)?

6. Can one establish that unlimited capacity for simple perceptual categories (e.g. letter vs digit) is not due to simple features?

7. Can one establish the minimum time for attention switching and hence for serial processing?

8. What are the relative roles of response selection, response production, and preparation in the PRP effect?

9. Can one extend the generality of the interpretation of PRP paradigm by developing related paradigms?

10. Can one clarify the capacity limits of short- and long-term memory encoding and retrieval?

11. Can one develop a more detailed theory of attention and learning based upon restructuring?