The basic cognitive processes of learning

A discussion paper by

Lesley Mackenzie-Robb
Vantaggio Ltd
www.vantaggio-learn.com
April 2005

This paper discusses the basic human cognitive processes that facilitate learning in adults, considering in turn, the processes involved with speech comprehension, written word recognition and visual perception. An understanding of these cognitive processes is fundamental to good pedagogic design for eLearning.

Speech comprehension has been the subject of considerable debate and study. As Samuel (1996) pointed out, theories are either based on an autonomous approach (e.g., Massaro, 1989, as cited Samuel, 1996) in which the acoustic signal contains all of the information necessary to comprehend meaning, or an interactive approach (e.g., Elman and McClelland, 1986; Marslen-Wilson and Clifton, 1980; Shillcock, 1990, as cited Samuel, 1996) in which knowledge and expectations influence meaning in an interactive activation of top down and bottom up processing. Many studies appear to support the notion of speech comprehension as being an interactive and constructive process (e.,g., Warren and Warren, Cutler et al (1987) and McGurk and MacDonald (1976), as cited Eysenck, 2000). There are, however, problems with the stability and predictability of empirical studies, in what Samuel (1996) describes as the "on-again, off-again" nature of lexical effects. As Samuel (1996) argues, some phonemic restoration studies have provided support for the top down effects of lexical representations having an effect on phonemic processing, but others have not.

The process of speech comprehension is vastly complicated by the physical manifestation of phoneme strings as often ambiguous, over-lapping and non-invariant. Most studies have focused on determining to what extent lexical and contextual knowledge, and expectations, facilitate speech comprehension, and the levels of interaction (or not) of bottom up and top down processing. Research methods have included Cross Model Priming tasks (Shillcock, 1990), word monitoring tasks (Marslen-Wilson and Tyler, 1980) and lexical decision tasks (Marslen-Wilson et al, 1996).

Samuel (1996)'s empirical study involved a number of experiments undertaken in laboratory conditions in which the lexical influences on phonemic restoration were tested using words and pseudo words. In the first two tests, the task for participants was to determine if the missing phoneme in the (pseudo) words in each case had been replaced entirely or superimposed with an extraneous noise. This was a variation on an earlier study by Warren and Warren (1970, as cited, Eysenck, 2000) in which was shown that participants would "fill in" the missing phoneme based on lexical context of the sentence, and often were unaware of doing so. Subsequent tests were designed to specifically address and assess the occurrence of lexical effects. The ultimate aim of Samuel (1996)'s study was to derive a set of conditions which would reliably lead to the observation of any lexical influence on phonemic perception, the issue being that inconsistencies in past studies had led to interpretive problems allowing both autonomous and interactive theorists to use empirical results to support their arguments. Samuel's (1996)

conclusion was that top down processing, in the form of lexical effects, can be shown to be present in speech comprehension, although they are "fragile" as they are not always present in all conditions, but none-the-less real.

Significantly, Samuel (1996) highlighted the important role of additive or replacement noise-type, phoneme class and the position of the target phoneme in the ability to observe lexical effects on phoneme restoration. It would logical to determine if these conditions remained reliable in the prediction of such effects in native speakers of other languages – such as Thai or Japanese – where there is a different phonemic construction, as a next step.

The debate now is to what extent these processes interact simultaneously (Marslen-Wilson et al, 1980, 1990, 1994, as cited, Eysenck, 2000), and the level of that interaction, and that this may be resolved by taking the view that bottom up processing is more at play where conditions – and therefore ability to perceive – are good, but that top down processing becomes more important when conditions are poor.

Similarly, there is diversity of opinion concerning cognitive processing in written word recognition. Two of the main debates centre on whether it is an automatic process and also whether word recognition is a purely lexical process as proposed by McClelland and Rumelhart (1981, as cited, Eysenck, 2000)'s Interactive Activation Model, or that it also necessarily involves phonological coding, a notion supported by Frost (1998, as cited, Eysenck, 2000). There are also fundamental issues with the manner in which research studies are carried out: many methods are "manufactured" and often entail "unnatural" activities. The difficulty is that reading is something that is carried out introspectively and sub-vocally, so gaining an accurate measurement of speed and accuracy is problematic with empirical methods.

A number of studies have been used to demonstrate that word recognition is a "fairly" automatic process (Chessman and Merikle, 1984; Marcel, 1983; Allport, 1986; Rayner et al, 1981, as cited, Eysenck, 2000), but it cannot be said to be wholly automatic as studies show that it does involve some processing, although not much. Rayner et al's (1981, as cited, Eysenck, 2000) Moving Window Reading Task used an infra-red beam to accurately measure participant's eye movements whilst reading controlled text from a screen, from which they drew the conclusion that as people don't spend a lot of time reading words (typically 200 – 250 milliseconds per fixation), then not much processing is involved. This, however, assumes a direct correlation and dependency between time and effort, which may not necessarily be the case.

Having said that, an automatic process is not necessarily one that excludes the involvement of both bottom up and top down processing. A study carried out by Jordan, Thomas and Scott-Brown (1999, as cited, Eysenck, 2000) explored the Illusory Letter Phenomenon, in a follow on from an earlier study by Reicher (1969) which established the Word Superiority Effect. This effect is almost identical to the Phonemic Restoration Effect discussed earlier. In this case, participants were shown words on screens placed some 3 metres distant. All of the words (the independent variables) consisted of four letters. Some of the words consisted of intact outer letters surrounding blurred internal "non-letters". Participants were asked to read the words, which they were able to do, including those with ambiguous letters. The conclusion drawn is that word recognition is an active, constructive process in which people interpret the shape and length of the words rather than solely their individual component letters. The influential McClelland and Rumelhart (1981, as cited, Eysenck, 2000) Interactive Animation Model predicts this effect and can be used to explain it as the means by which bottom up processing is used to perceive the features of the word and its letters. Top down processing is used to activate possible representations of the word in an iterative process that interacts top to bottom until a stored word is retrieved which accurately matches the features presented.

In both the Jordan et al (1999) study and the McClelland et al (1981) theory of word recognition, there are issues and limitations. Both only apply to four-letter words, and neither adequately explains the phenomenon of recognising words when they are misspelled. There is also the debate concerning pre-conscious processing of

words, and at what point processing becomes conscious. What can be offered is that word recognition is a relatively effortless, unavoidable and therefore fairly automatic process, and that it does involve both bottom up and top down processing, but the precise nature of this has not been demonstrated.

Opinions and theories concerning visual perception are equally diverse: the Constructivists believe that top down or concept driven processing is fundamental to perception, and the Direct theorists subscribe to the notion that perception is solely reliant on bottom up or data driven processing in which information presented in the external environment is sufficient without the need for inferences drawn from existing knowledge and expectations. Neisser (1976, as cited, Eysenck, 2000) offered a synthesis of these approaches in his "Perceptual Model". What is interesting is that all three concepts were developed virtually simultaneously. As with other areas of study, there are also considerable difficulties in conducting empirical studies in this area, as, inevitably, such research expands the field from the study of visual perception to the study of recognition.

Roth (1986, as cited, Eysenck, 2000) clarified the process of visual perceptions as being "the means by which information acquired via the sense organs is transformed into experiences of objects, events, sounds, tastes, etc." Dodwell (1995, as cited, Gross, 2001) expressed its complexity: "…To perceive seems effortless. To understand perception is never the less a great challenge." In a comparison between Direct and Indirect approaches to visual processing, it becomes evident that the Constructivists are investigating the processes from a perspective of "seeing as" or recognition, whereas the Direct theorists focus on the process of seeing only. As the Neisser (1976) model shows, there is a need to acknowledge that there is validity in both the Constructivist and Direct theories.

Vecera and Farah (1997) undertook a study of image segmentation processes to determine whether this is a bottom-up data driven or a top-down concept driven process. They raise an interesting paradox: the goal of image segmentation is to visually group the features of an object in order to recognize it, therefore how can pre-stored information about the object be used to fuel this process when it hasn't been recognized yet? Vecera et al (1997) conducted a number of experiments using letter shapes and non-letter shapes containing "x"s. The first of these used overlapping letters - upright, rotated and non-letter shapes – where the task was to determine which shape contained the "x"s. By using ambiguous letters, the aim was to manipulate letters' familiarity and therefore the availability of internally stored object related information. The notion was that if only bottom up processing is involved, there should be no difference in the speed of perception between the upright letters and the others, whereas if top down processing is involved, there should be marked differences. This latter was shown to the case, with response times significantly faster with normal letters, than with rotated or non-letters, and this is consistent with the predictions of an interactive model. The experiment did raise further questions in respect of the response times to the rotated letters, and why these were not significantly different from those to non-letters.

Vecere et al (1997)'s final experiment in the series is the more convincing one. The aim was to determine if top down processing could over-ride bottom up grouping heuristics by tasking participants to identify if "x"s appeared in the same region of a single shape, or on both shapes comprising the block. Various combinations were used in each block, including an upright letter shape combined with a non-intact letter shape. In some cases, the combination was designed to look like an emergent letter. The important finding here was that "x"s appearing on different regions were grouped together where those regions could be construed to be an upright letter. In this case, top down processing is shown to over-ride bottom up features, which is wholly inconsistent with a purely bottom up model.

The experiments can be used to conclude that image segmentation in visual perception is an interactive process in that it uses both top down and bottom up processing, at least part of the time. In this case, good familiarity enabled stored information to be used, whereas unfamiliarity prevented the use of lexical information. Having said that, even Vecera et al (1997) acknowledged the limitations of their experiments (e.g., stimuli presented for short

periods of time, the use of transparent letters). To what extent such results could be said to be universal is debatable.

These studies, and other empirical studies do not offer a full and conclusive explanation of the processes involved and, moreover, they are almost always undertaken in unnatural environments, and involve unnatural actions on the part of the participants. As seen in the Vecera and Farah (1997) study, they often raise more questions and issues than were started with. There are probably more theories in visual perception than any other area of study, but none has offered a complete solution, and few have been the subject of detailed empirical study. So, although it can be suggested that bottom up and top down processing are involved in visual perception, firm and supported knowledge about how visual processing systems work is more the result of physiological research than cognitive research.

In conclusion, empirical studies in these three areas can be said to provide indicative support to the notion of bottom up and top down processing being important factors, but this is far from conclusive, and understanding of these complex areas is far from complete.

Bibliography

Velmans, M; . When Perception Becomes Conscious, British Journal of Psychology, 1999, 90, 543-566

Samuel, A.G.; Does Lexical Information Influence the Perceptual Restoration of Phonemes?, Journal of Experimental Psychology, March 1996, Volume 125 (1), 28-51

Samuel, A.; Phoneme Restoration, Language and Cognitive Processes, 1996, II (6), 647-653

Eysenck, M.W., Keane, M.T.; Cognitive Psychology, 4th Edition

Gross, R.; Psychology: the Science of Mind and Behaviour, 4th Edition

Vecerera, S.P.and Farah, M.J.; Is Visual image segmentation a bottom-up or an interactive process?, Perception & Psychophysics, 1997, 59 (8), 1280-1296