An alternative account for the effects of age of acquisition[1]

Abstract

This chapter is organised to present a brief introduction of age of acquisition (henceforth AoA) and why it has taken so long for acceptability. This is followed by four accounts that attempt to explain AoA effects, including my own working hypothesis[2]. From these accounts, predictions derived are tested in a battery of experiments on celebrities' faces and printed names. Final experiments demonstrate the importance of '*affect'* on naming celebrity' faces.

Key words: AoA; SSPS; Affect.

Introduction

An extensive literature reports that high-frequency words are processed more quickly than are low-frequency words (see Monsell, Doyle & Haggard, 1989 for review). For example, people are faster to decide that a high-frequency word constitutes an English word than a low-frequency word in a lexical decision task (e.g. Scarborough, Cortese & Scarborough, 1977). People are faster to name pictures with high-frequency names than low-frequency names (e.g. Oldfield & Wingfield, 1964, 1965; Jescheniak & Levelt, 1994). In such studies, measures of word frequency were derived from published corpora, e.g. Kuèera and Francis (1967) and CELEX (Baayen, Piepenbrock & van Rijn, 1993). More recently, a number of studies demonstrate that AoA is an important predictor of the speed of lexical processing, i.e. people are faster to name pictures with names learned early in life than pictures with late-acquired names. This effect has been demonstrated in English (Morrison, Ellis & Quinlan, 1992); in Spanish (Cuetos, Ellis & Alvarez, 1999) and in written French object naming (e.g. Bonin, Fayol, & Chalard, 2001; Bonin, Chalard, Méot & Fayol, 2002). An AoA effect has also been demonstrated in written word production in French (Bonin *et al.*, 2001; Bonin *et al.*, 2002); on English and Dutch word reading and lexical decision tasks LDTs (e.g. Morrison & Ellis, 1995; Bonin, Chalard, Méot, & Fayol, 2001; Brysbaert, Lange &

[1] Acknowledgements. This research was supported by a UK Economic and Social Science Research Council grant (R000 429624208). I would like to thank many researchers for their patience and time in discussing the ideas presented here. Tim Valentine Goldsmiths College, University of London; Andy Ellis and his research team the University of York, Patrick Bonnin and his research team at the University of Clermont-Ferrand, France, Debbie Roberson University of Essex, Judy Turner, University of Reading. Vickie Bruce, University of Edinburgh. I would also like to thank Kelly Brown, Zoe Knight, Emma Rickaby, Kim Shuntara and Katherine Worral for collecting the data for the reported experiments.

[2] Some of this work was presented in the paper Moore, V. & Valentine, T. (2001) to the ESCOPs and to the HRD of Dr Patrick Bonin (2002).

Van Wijnendaele, 2000) and on an auditory LDTs in English (Turner, Valentine & Ellis, 1998). The AoA effects are not confined to lexical processing. The age at which a celebrity was first encountered was demonstrated to affect the time taken to name faces, read names aloud and to decide whether a face or name is familiar or not (Moore & Valentine, 1998; 1999). These AoA studies controlled for word frequency or rated familiarity (that is the cumulative frequency with celebrities) of early- and late-acquired items. However, many past and contemporary studies reporting the effects of word frequency fail to control for AoA. Some researchers argue that AoA effects are 'interchangeable' with word frequency effects (Seidenberg, Peterson, MacDonald & Plaut, 1996) or re-interpreted the AoA effects as 'trajectory frequency' "how experience with a word is distributed over time" (Zevin & Seidenberg, 2002, page 10). This view is seductive because AoA and word frequency are highly correlated. However, there is mounting empirical evidence of AoA effects in domains where such correlations are not so strong. For example, AoA effects were established in naming and recognising faces and names of celebrities (Moore & Valentine, 1998). There are also empirical studies that carefully disentangled these two variables and demonstrated AoA effects in a task (i.e. auditory LDT) that does not show word frequency effects (e.g. Turner *et al.*, 1998).

The problem of separating word frequency and AoA effects are compounded because some 'objective' measures of word frequency and AoA were taken from the same source (e.g. Rinsland, 1945). These corpora were designed to advise teachers when children should learn particular words, and it was assumed that because children learn these words early in life, the words would 'naturally' acquire a higher 'cumulative frequency'. However, there has been a consistent failure to find an independent *empirical* influence of cumulative frequency on lexical processing tasks (Carroll & White, 1973a; Gilhooly & Gilhooly, 1979; Gilhooly, 1984; Morrison, Chappell & Ellis, 1997; Morrison, Hirsh, Chappell & Ellis, 2002).

The Phonological Completeness Hypothesis: The locus of AoA effects is subject to a continued debate in the literature. The earliest account for AoA effects is the phonological completeness hypothesis (Brown & Watson, 1987) that attributed these effects to the phonological representation of words. Accordingly, this hypothesis requires a functionally different mechanism to occur because the phonological representation of early-acquired words are stored in a more *complete* form than representations of late-acquired words (requiring re-assembly for production). Consistent with this view Morrison *et al*. (1992)

located AoA effects in object naming at the speech output lexicon, supported by a null effect on semantic classification and a delayed naming tasks. Morrison and Ellis (1995) report significant AoA effects in visual LDTs, which they attributed to an automatic activation of phonology in the output lexicon. Gerhand and Barry (1999) replicated this effect reporting the use of orthographically illegal non-words and pseudohomophones as non-words failing to extinguish AoA effects, also the use of articulatory suppression as a secondary task (repetition of a nursery rhyme or the word "the") while making decisions failed to remove significant AoA effects. Notwithstanding this finding, Gerhand and Barry also attributed the AoA effect to an automatic activation of phonology from seeing the word.

The locus of AoA effects at speech output implies that these effects arise from learning during a 'critical period' of language development, when the representations underpinning speech are acquired. Theories of infant language development provide support for this locus. A developmental view of language specificity proposes infants to be innately equipped to process tone, stress, vowel length, etc. of the world's languages and they become attuned to phonemic contrasts in their linguistic environment during the first year (Werker, 1994). Once established, these are used to discover regularities in speech, where infants by nine months, show a 'preference' for listening to words rather than non-words (Juscyk, Cutler & Redanz, 1993). Infants show a 'preference' for listening to phoneme structures conforming to their own language (Juscyk & Aslin, 1995) implying language regularities are used to hypothesise word boundaries in speech streams. Furthermore, infants use the rhythm-type to decide which segmentation unit to use for further speech analysis (Kokelaar, 2000). Infants can learn novel words from as few as nine presentations, suggesting a powerful learning mechanism for forming object-label associations exists (Woodward, Markman & Fitzsimmons, 1994). By around eighteen months a 'spurt' of language comprehension and often production (Goldfield & Reznick, 1992) suggests the triggering of a new principle of organisation into the child's understanding of the object-to-label relationship. This is consistent with a 'critical' period of language development forming an AoA locus at phonological output and phonological *input* (supported by AoA effects on auditory LDTs Turner *et al.,* 1998). This may also explain the absence of AoA effects on semantic tasks (Morrison *et al.,* 1992) because, the stress occurs for associations between an object's appearance and its name. Despite this evidence, a single locus for AoA is not universally supported. Yamazaki, Ellis, Morrison and Lambon Ralph (1997) report that the reading speed of Japanese Kanji characters was affected by the age at which words entered the *spoken* vocabulary *and* the age the *written* characters were learned. Yamazaki *et al.* argue that AoA affects the quality of lexical representations in the speech output *and* visual input lexicons, requiring at least two loci of AoA.

Consistent with these findings, Moore and Valentine (1998) reported early-acquired celebrity' faces and names were named faster than late-acquired faces and names. These effects were interpreted using a functional model of face recognition (e.g. Bruce & Young, 1986) that evolved anabgous with the logogen model of word and object recognition. The debate concerning 'special processing for face stimuli' is not relevant here, as the empirical data demonstrate AoA effects are found in word, object and person processing tasks. Therefore, it is reasonable to 'explain' such effects using a generic functional model applicable to all stimuli. However, it is important that the AoA effects reported by Moore and Valentine (1998, etc) do *not* represent a model of developing face recognition *per se*. Infants from a few days old are already learning the surrounding faces. The celebrities used in our experiments were acquired between 6 and 12 years of age therefore, these AoA effects require a rather different explanation than a phonological locus, or mechanism. Le Grand, Mondloch, Maurer and Brent (2001) demonstrated that deprivation of patterned visual input from birth until six months of age result in permanent face processing deficits. They argue that the early visual input is necessary for normal neural architecture to develop. They argue that early information is responsible for setting up the neural architecture that will come to specialise in expert processing of faces over the next 10 to 12 years. Therefore, the AoA effects reported in our experiments represent something other than 'initial' face processing, rather initial processing of famous people as a set of novel stimuli. Therefore, AoA effects in person processing challenges a single locus of AoA and the notion of a critical period. Moore and Valentine (1999) report that AoA affected the speed in a familiarity decision task (FDT) whether a famous face was familiar or not. Faces are notoriously difficult to name (see Valentine, Brennen & Brédart (1996) repetition priming studies demonstrate that names are not automatically activated during a face FDT (Valentine, Hollis & Moore, 1998). These data are difficult to reconcile with a 'phonological' locus because no verbal response was required. Furthermore, famous faces and names are acquired later in life than typical object names, suggesting that AoA influences cognitive processing beyond a 'critical period' of language acquisition.

Cumulative frequency has been proposed to account for the AoA effects (e.g. see Caroll & White, 1973a and Gilhooly, 1984). A recent account being an instance-based model in which the reaction time to a stimulus is given by a function that includes a negative power of the number of instances in memory and a positive power of the time since last exposure to the stimulus (Lewis, 1999). Lewis argues that this function accounts for AoA effects found in recognition and naming tasks, and that AoA effects are explicable as cumulative frequency on retrieval from the semantic system. This approach challenges those who suggested a speech output locus for AoA effects. Lewis claimed to have demonstrated cumulative frequency effect when celebrities were classified according to which of two British TV 'soap operas' they appeared. The task was to classify celebrities as belonging to either Coronation Street or East Enders. However, the data is confounded in a number of ways:- the measure of AoA was an estimated (by the experimenter) number of instances were in participants' memory (what Moore & Valentine, 1998, etc., refer to as 'familiarity' or 'frequency of encounter'). No measures of, or control for, rated AoA, familiarity or facial distinctiveness, included as a matter of course in other studies. An assumption that the actors were known uniquely for their one soap-opera role, although for some, fame preceded the program (e.g. Mike Reid). Some celebrities appeared in other contemporary top-rated British TV series and plays (e.g. Sarah Lancashire). The experimental stimuli represented close semantic associates, where the reaction times (RTs) might have been affected by semantic or associative priming (Bruce & Valentine, 1985). A raised semantic activation may have occurred because of the large number of celebrities derived from the same category (Sergent & Poncet, 1990). Finally, the speed of the classifications[3] the mean classification RT's (1300ms, N = 9) is not comparable with mean RTs taken from larger samples. For example in face FDT (mean RT = 662ms.; N=96); occupation classification tasks (mean RT = 898ms.; N = 96)[4]; naming faces (mean RT= 1490ms.; N = 800) (Moore, 1998). Despite these misgivings, the data reveals that the classifications times are faster for familial than none-familial celebrities. This was confirmed in a reanalysis of the data between familial and non-familial celebrities: East Enders ($F(1,92)=30$, $p<.0001$); Coronation Street ($F(1,84)=64$, $p<.0001$)[5] see Moore, Valentine and Turner (1999) for a critique.

The Neural Plasticity Hypothesis: Many connectionist models of human cognition were designed to account for the effects of word frequency. However, until recently few could account for AoA effects. Seidenberg and McClelland's (1989) used backward error-propagation (backprop) to learn mappings between orthography and phonology, and could account for a wide range of effects in normal reading e.g. it simulated the effects of word frequency on word naming latency. However, they were unable of simulate empirical AoA effects. Until recently it was not clear how Seidenberg and McClelland's model could account for the persistent empirical AoA effects because backprop was employed for the network to learn

[3] The Burton, Bruce and Johnson's (1990) functional model of person processing suggest that processing speed, as determined by RT's reflect the processing hierarchy. Accordingly, a familiarity decision made to a face will be fast because access to the PIN is required. Semantic classifications will be made more slowly, because of the extra stage of processing. Face naming would be the slowest as the final stage of processing.

[4] This was a very contrived task, participants decided whether each celebrity was famous or not for appearing as a chat-show-host (Moore, 1998) and therefore does not reflect an on-line task. Thanks to Vicki Bruce for this interpretation.

[5] This effect was robust analysing raw data, logarithmic transformations and reciprocal transformations.

distributed representations. In previous attempts to model AoA on backprop networks the introduction of a second set of patterns served to wipeout previous (early-acquired) learning (i.e. catastrophic interference). Humans learn sequentially, where learning a new pattern of behaviour, or word, would not extinguish previous learning, but should facilitate new learning.

Ellis and Lambon Ralph (2000) argued that AoA effects are a 'natural property' of backprop networks. This is in terms of the behaviour of a distributed memory network trained by backpropagation (these networks are widely used to simulate aspects of adult lexical processing). Their network was first trained on one set of 'early-patterns'. Then later-patterns were added, with all of the patterns being trained in a cumulative and interleaved fashion. The performance of the network remained superior on early-, compared with late-patterns, even after extensive further training. There was a robust advantage for the early-acquired set over several simulations even with number of learning trials equated. Analysis of the hidden units suggested that the AoA effects were caused by a gradual reduction in the network's plasticity, causing a consequent failure of the network to differentiate late-items as effectively as early-patterns. Therefore, early-encountered patterns played a prominent role in the organisation of the representational inputs. Monaghan and Ellis (2002) argue that AoA is modulated by the similarity, between early- and late-patterns. When low correlations occurred, the network structure (built up during training on early-patterns) did not help the assimilation of the later-learned patterns. Therefore, the network failed to attain representations comparable to those of early-patterns. However, when the correlation between early- and late-patterns was high, the late-patterns took advantage of the network structure established by the early-patterns. This explanation makes intuitive sense of how humans learn and expand their information base.

The neural plasticity hypothesis is a valuable tool for understanding AoA effects on lexical processing, however, on closer examination some limitations appear. The original model describes early-learned patterns reducing neural plasticity with less room for later-learned patterns. This does not appear to reflect how children and adults learn and continue to learn throughout life. This predicts the diminution of learning with age and while it is generally known that 'aged' people do have problems learning certain types of new information (e.g. word lists) however, it cannot explain the expertise involved when people change profession and/or continue to acquire specialised knowledge throughout life. The later adaptation of correlational facilitation reflects patterns of human learning with early-learned patterns facilitating the acquisition of later patterns. However, it must be remembered that backprop networks are neurological *implausible* (O'Reilly & Mankato, 2000), therefore these models cannot explain the how and especially the why *people* learn certain types of information over other types of information at different stages of life. While networks learn patterns of information this requires stimuli to be presented by an experimenter or program as 'input' of material to be learned. This cannot explain the dynamic and interactive process of human learning and especially the changes from novice to expert throughout life. For example recent findings demonstrating navigation-related structural changes in the hippocampi of taxi drivers (Maguire, Gadian, Johnsrude, Good, Ashburner, Frackowiak, & Frith, 2000). Furthermore, the idea of reducing neural plasticity cannot explain the empirical AoA effects on the processing of celebrity' faces and names even with the additional correlation account.

The developmental theories of language acquisition served to demonstrate how the phonological completeness account for AoA effects could have arisen during the critical period. An alternative argument may be that these effects arise from the *order* in which information is acquired. I argue above that a critical period of language development cannot account for the AoA effects in person processing. It is possible that an explanation of *temporal order* of acquisition could in some way account for these effects. If so, this would be supported by data from the neuropsychological literature where a temporal gradient of memory loss is an intrinsic diagnostic tool

Preserved memory in organic amnesia: Many studies report deficits in organic amnesia, retrograde (pre-morbid memory) and anterograde amnesia (acquiring new memories). These temporal order of preserved memory have not yet been linked with AoA *per se*, however, it is demonstrated in some patients with severe anterograde amnesia. For example, SS, a 65 year-old man with organic amnesia, evinced an effect of order of acquisition (Oaoa) (Verfaellie, Croce & Milberg, 1995). SS was tested on words and concepts, for which entry into the English language was dated into hemi-decades. Pre-morbid items entered the language between 1920 to 1970, post morbid items between 1971 and 1990. SS recalled and recognised the meaning of novel words acquired between 1970 and 1980 significantly better than words acquired in the 1980s, although both had entered the vocabulary *after* the onset of his amnesia. Shallice and Kartsounis (1993) report WK, a 56-year-old man with a person-naming deficit. WK could name familiar personalities, famous 20 years ago, e.g. Harold Wilson (British prime minister 1964 to 1976), but not Margaret Thatcher (the contemporary British prime minister). He named historical but not contemporary media personalities although the Oaoa effect generalised to naming from definitions of new words (from the past 20 years). For example WK knew that video was '*a device to record TV programmes so that one can see them later…* '. However, he could not provide the modern definition of 'aids'. Baddeley (2000) reported a densely amnesic patient who continued to play bridge successfully, this was explained by the 'working memory' model's 'episodic buffer' which Baddeley argued was preserved in this case as a temporary store capable of holding complex information and involved in manipulating it to utilise it over time. Clearly, the episodic buffer could be responsible for holding complex information and even manipulating that information. However, what is unclear is *how* or *why* an isolated island of memory continued to exist for this specific interest alone.

It is possible that the developmental, neuropsychological and face naming literature may offer converging support for an effect of a *temporal Oaoa*. To understand the mechanism(s) giving rise to age or order effects it is necessary to explore these effects in a domain where items are acquired after language development. Therefore, processing famous faces and names provided an ideal domain to determine whether these effects were indeed specific to language acquisition. Recognition of famous faces and names are particularly suitable because many current theories of face and name processing were developed by analogy to theories of object and visual word recognition using a similar hierarchy of representations for naming both faces and objects (see Tarr & Cheng, 2003, for review). Moreover, it is important to recognise that the hypotheses presented above cannot account for the empirical face and name processing data. The model presented below is consistent with *a* locus for AoA effects to be at the phonological output level; it is consistent with the idea of neural plasticity; the importance of frequent exposure, it may also explain the temporal order and preserved isolated islands of memory.

Set up of Specialised Processing System (SSPS) Hypothesis is necessarily complex as it evolved from a variety psychological areas. The simple take-home-message is that a sub-threshold level of *affect* is created when any new order of stimuli are encountered. When sufficient encounters have occurred the new SSPS system will process that type of information in a more shorthand way. The unique feature of the SSPS hypothesis lies with the incorporation of *affect* as the heart of cognitive processing. This is explained from its beginnings, rooted in preserved memory in organic amnesia, some neuropsychological explanations of arousal/emotion/affect I will demonstrate how a physiological sub-threshold level of arousal *primes* motor movements and thus may account for the AoA effects on spoken, button-push and written responses.

The SSPS hypothesis began while researching preserved implicit memory in organic amnesia (Mayes, Poole [Moore] & Gooding, 1991). Families reported their 'amnesic' to have a specific area of memory that seemed to be spared e.g. one woman's pre-morbid interest in pop-music continued as she learned parts of new tunes *post morbidity*. A boy M. had mimicked celebrities from childhood, he explicitly recalled a meeting with a famous UK celebrity, repeating what the celebrity had said to him, mimicking the voice, intonations, etc. These foreshadow the bridge-playing patient reported by Baddeley (2000). However, the most striking event occurred while testing JS (Mayes, Downes, McDonald, Poole, [Moore], Rooke, Sagar, & Meudell, 1994). JS was densely amnesic; he had attended a day centre since his amnesia became very severe three years previous. While testing JS, the fire alarm sounded. I suggested we leave, JS said, "No, the fire alarm always goes off at three o'clock on Wednesday afternoons" Therefore, JS had *explicitly* recalled this information, although according to the contemporary literature this was not possible as the acquisition of implicit memories was under debate at that time (Mayes, *et al.,* 1991). Since then the *why* and the *how* some people retained this island of preserved memory has been facinating. The first two people had a history of interest in specific areas. However, the only explanation for JS was that salience (attention grabbing response) of "fire alarm" suggested previously JS had reacted to hearing the fire alarm, until his arousal system had 'learned' not to react. This suggested that some theoretical level was missing from the cognitive and neuropsychological literature and the following hypothesis began. As such this is not a thesis on *age* or *order* of acquisition *per se* but an attempt to integrate several psychological areas for researchers to develop a more unified theory of cognition, than currently exists. In this chapter I, construct a working hypothesis using AoA from built a framework is built, incorporating convergent evidence combined a novel way to offer the Set up of a Specialised Processing System (SSPS) for consideration of researchers in the disparate fields.

The term AoA is employed here for continuity with the concept; however, it does *not* apply to a critical period of infant development, but refers to the development along a chronological dimension. I argue that AoA effects are *caused by the first exposure to novel exemplars of information at any age* which allows further, similar exemplars to be processed in a more automatised fashion[6]. I suggest that AoA reflects a temporal *order* of acquisition, that is to say the first presentation of novel items in a specific stimulus domain would be the initial exemplars of new types of information to stimulate a physiological, orienting response[7], which would initiate the set up of a specialised processing system (SSPS). In this way,

[6] Note Langer, 2000, argues that such automa tisity restricts diverse interpretations and encourages simplified learning, also see Killcross, 1999, but is consistent with Ellis and various colleagues view of neural plasticity.

[7] Such a response may be that reported as an attentional mechanism. For example, pre-pulse inhibition occurs when a weak non-startle stimulus is presented (30-500ms) before a startle probe and serves to reduce the effect of the probe

AoA should affect both perceptual and output motor movement on processing task for which specialised representations have been established (Moore, 2000; Moore & Valentine, 2001).

We were the first to demonstrate AoA affecting perceptual tasks, that celebrity' faces and names rated as early-acquired were recognised as famous faster than late-acquired celebrities in a FDT (Moore & Valentine, 1999). We recently demonstrated this on object perception (Moore, Smith Spark & Valentine, in press) with faster classification times to pictures of early-acquired objects (as 'real objects' rather than none-objects) than to late-acquired objects. These tasks were designed as analogous with lexical decision tasks (LDT) for words and are important because until very recently the effect of AoA on LDT was argued to be the automatic activation of phonological output, but this cannot explain these effects on objects or celebrity' faces. The SSPS hypothesis alone proposes a perceptual locus for AoA effects because new exemplars of information are the activation creating the basic restructuring process. This hypothesis evolved from the question of *how* and the question *why* the human brain should or would develop 'recognition units' or an independent pattern of activation for specific exemplars of information (the SSPS can be interpreted by modular or parallel processing explanations). I argue that a raised level, affect/arousal/emotion and/or motivation is required to 'kick-start' this process. For the purposes of brevity, henceforth this is referred to as *affect* alone. In line with this proposal, Anderson and Phelps (2001) investigated the function of the amygdala in enhancing the perception of emotionally significant stimuli and argued that their results reveal a neural substrate for affective influences on perception. They suggest that this indicates a similar neural mechanism may underlie the affective modulation of both recollective and perceptual experiences. A startle response with the consequent physiological arousal is a defence mechanism causing an immediate interruption of ongoing behaviour and the release of neurochemicals (Lang, 1995). The startle response is an automatic (Yantis, 1995) attention orienting system, driven by stimulus properties that in turn, cause a series of rapid flexor movements e.g. eye blink. I argue that a lower, sub-threshold level of *affect* that would attenuate the physiological startle but would orientate the individual's attention. Pinpointing this level of *affect* is not a simple matter, e.g. exposure to new exemplars of stimuli causes some raised level of *affect* – this in turn would warrant the increased level of activation required to initiate heightened vigilance for similar exemplars. That is an attentional interaction of bottom up initiating the top-down vigilance. There is support for this sub-threshold level of *affect* derived from prepulse inhibition (PPI) research where it has been noted that the full startle reflex can be

---

stimulus (Lang, Bradley & Cuthbert, 1997). They argue that the operation of pre-attentive processes serve to increase vigilance for that type of information.

reduced by a prepulse attentional mechanism. Pre-pulse inhibition occurs when a weak non-startle stimulus is presented shortly (30-500ms) before a startle probe and reduces the effects of the startle probe. Bradley, Cuthbert and Lang (1993) report that affective stimuli can act as PPI to reduce the startle reaction. When unpleasant and pleasant pictures were presented shortly before startle stimuli, the magnitude of the eye-blink was reduced compared to the response for neutral pictures. Lang *et al.*, (1997) report that PPI caused by affective pictures reflects the operation of PPI and argue this is due to affective pictures reflecting the operation of *pre-attentive* processes mediating increased vigilance. This is consistent with Ohman and Soares (1994) reports where animal phobics evinced elevated skin conductance to pictures (e.g. snakes) with phobic content, non-phobics did not. Furthermore, this effect was robust with presentation times below conscious recognition (between 30 and 500ms, Graham, 1980; and 130ms, Tipples & Sharma, 2000). Ohman (1996) argued that such findings support an evolutionary mechanism for automatically detecting emotional salience. I argue that this is equivalent with the affective response to new exemplars; a mechanism capable of attenuating the startle response can explain why individuals orientate *away* from something 'unpleasant' and *towards* something 'pleasant' in the array of available information. That is to say, a bottom-up-top-down interactive process of scanning the environment based on what has gone before assuming that perceptual and sensory channels monitor the environment with a bi-directional flow of input to and from the affective system for unknown, to be preferred or to be aversive stimuli (by updating the semantic system). For example, a gunshot in the countryside may not generate a startle response, so a sudden bang is not always startling. A startle with accompanying physiobgical arousal may be caused by a stair-board creaking late at night. Therefore, it is that to which one is *unfamiliar* that initiates the orienting attentive response. This 'pre-startle' raise in a physiological arousal may serve to *prime* an appropriate motor response, i.e. speech or movement. This would also explain why AoA did not affect delayed naming (e.g. Morrison & Ellis, 1995). Evidence for the 'prime' of motor movement may be derived from the empirical AoA effects reported on written tasks (e.g. Bonin *et al.*, 2001; Bonin *et al.*, 2002).

To summarise, the SSPS hypothesis evolved from questions of *how* and *why* should or would the human brain develop 'recognition units' (or independent pattern of activation) for specific exemplars of information. I argue that a level of *affect* must be an intrinsic precursor to any such development. Through the SSPS hypothesis, I propose that new exemplars of information (if perceived in frequently within an individual's environment) will initiate a fundamental restructuring of perceptual processing. This will, in effect, peg out the parameters for the specific type of processing required for future similar, exemplars. There exists a body of converging evidence to support this argument. Simon (1967) argued that general theories of thinking and problem solving must incorporate the influences of emotion/*affect* and saw this as a powerful motivator, influencing perception, cognition, coping and creativity in important ways. The neurosciences have indicated a pivotal role for emotion in attention, planning, reasoning, learning, memory and decision-making (e.g. Izzard, 1993). It has been argued that the demands of a neural system with finite resources required to operate in unpredictable and ever complex environments needs recourse to emotions for a decisive but flexible way of dealing with challenges both intelligently and efficiently. Indeed, Damasio (1996) argued that emotion is required for general heuristics to operate for perception and cognition.

Support for the emotion/*affect* in decision-making has been demonstrated by studies investigating the 'feeling of knowing', i.e. deciding whether someone is good or bad, likeable or not. When referring to people, is not always a decision made logically or rationally. Such valenced (i.e. negative or positive feeling) decisions are made without conscious awareness of any decision having been made. One does not always have access to stored memory representations when making decisions and we trust to our *experience* or *intuition*. Plus, we may know that we 'know' something but not be able to access the name or other information, however, this is combined with a 'feeling or gut reaction' that guides one towards or away from one or another name or concept. Conscious awareness may be apparent when the experience can labelled, e.g. a loud bang or creaking stair, and associated responses of increased heart rate, sweating, etc. may be labelled 'fear'. While consciousness is not required to experience this feeling, it may provide enhanced understanding and ability to re/act rapidly and on becoming aware that the cause of these 'feelings' can be avoided in by forward planning. Thus, secondary awareness may allow for many potential actions for future events.

Studies of human emotion have concentrated on finding universal patterns for interpreting the emotion in others, as in facial expressions, or expressing and feeling an emotion. In this way has been convenient for cognitive modellers to add another unit or link to a model and account for processing that specific behaviour. However, I argue that the power of *affect* is a primal requisite for creating the individual learning patterns. Furthermore, that AoA effects arise from a sub-startle reaction to novel exemplars that do not readily map onto previous knowledge. Furthermore, that such 'labelling' may be sub-vocal, it is where the semantic system has given a 'no info' message. The sense of unfamiliarity would cause the rise in *affect* and is where a new sub-system begins to develop which is the triggering of a new principle of organisation of the new exemplar category. This is pegging-out of the parameters to understand and incorporate such new information into one's semantic system. Therefore, the prediction is derived that the locus of AoA effects are peripheral to the semantic system, i.e. if representations already exist a feeling of familiarity should occur preventing the raised autonomic arousal, instead a mapping-on of new information relating to previous knowledge will occur. Therefore, the SSPS predicts that semantic memory would not be affected by AoA. This is consistent with the dual route of perceptual processing subserved by two communication systems – a rough recognition system that acts fast, e.g. danger detection, error prone, act first think later. Then a 'finer' pattern-type recognition, a slower and more precise system, which allows experience from the semantic system to influence the process currently underway (Damasio, 1996).

Damasio (1996; 1999) argued that there are certain features of stimuli in the world where a primary state (non-cognitively-generated) responds emotionally, i.e. a startle response which is primarily in and caused by the limbic system - innate responses. This activates a corresponding, cognitively generated, and later state of secondary emotions when systematic connections are identified between primary emotions and categories of objects and situations. These secondary emotions still activate the limbic system, but pre-frontal and somatosensory cortices are also involved. In particular, secondary emotions can be initiated simply by thought. Damasio reports a patient with damage to the frontal cortex, Elliot had intact primary emotions (startle); he understood that he 'should feel disturbed' to distressing pictures. However, he could not experience the actual disturbed 'feeling', and so evinced no physiological measures. Thus, although aware that, pre-traumatically, 'disturbing images' would have influenced his feelings, post-morbidly Elliot was unable to 'feel' them. Therefore, the cognitive thoughts that would normally activate or induce the emotion were dissociated from the sensation. Damasio argued for the requirement of limbic-cortico pathways to be intact for the secondary response to manifest the distinction between thinking "this is disturbing" and feeling "disturbed" can be dissociated.

The limbic system/s has a powerful influence on emotion, memory and attention and implicated with valence decisions and salience. Neural imaging studies reveal that the brain does not separate cortical and limbic processes, therefore, we should not assume that thinking and feeling are separated either. An example for a requirement for an *affective* component in learning was demonstrated (LeDoux, 1990) when rats were found to respond to an auditory signal even when their auditory cortex was completely removed. It was found that sections of the subcortical area appeared to receive the auditory signal *before* the auditory cortex. Lesions to the subcortical area (and more specifically the amygdala) eliminated the rat's ability to learn the 'fear tone' thus supporting a dual route of perceptual processing – quick and dirty direct to the limbic system and capable of highjacking cortical processing. The second cortical processing would be slower but more accurate. The neurological evidence supports the idea that emotions can highjack the mediating cognitive processes of the brain, such as when one jumps in fright and then realise there is no reason for fear.

It is possible that pre-natal infants are open to the effects of *affect*. During development, the foetus is active, its brain developing about 150 new cells per minute, with some areas almost fully developed at birth. This literature may lend support for the *affective* requirement to account for the effects of AoA or Oaoa. Measurements taken during the last trimester suggest the infant auditory cortex is developed almost to the level of the adult system (compared to the visual or other sub-systems). Hepper (1991) reported infants in utero kicked strenuously and evinced an accelerated heart-rate to the music of Bach. When the same music was played repeatedly the heart-rate and kicking declined, however, a change of music caused the kicking and heart-rate to increase again. in-In utero learning was reported by DeCasper and Spence (1986). A story was read aloud during the last trimester caused the post-natal infant to suck harder on a pacifier to get that story over another story. It was argued that the foetus picked up the rhythm, intonation, etc., patterns in speech. The infant may also pick up a sensation from the parent. In this way, this way a foetus can be seen as an active organism, learning and discriminating between the available stimuli because of familiarity.

At this point, the reader may argue that Ellis and Lambon Ralph's AoA simulations did not require *affect* or reference to emotions. Furthermore, the simulations of empirical and patient data were implemented without recourse to an *affect* component. Therefore, does it necessarily follow that learning requires recourse to a level of *affect*? I argue that this is, the case. Computer networks do not acquire information in the same way as humans because an operator or program provides and inputs the data. These networks do not passively 'acquire' representations of the presented patterns. The programmer presses a key thus the *affect* would be the connective energy transmitted through the process of inputting the 'to-be-learned-patterns'.

A general psychological rubric appears to argue that if humans perform 'such-&-such' a process, then it should be explicable in terms of a mathematic equation and/or to implement the process. The argument presented above suggests that the person or program providing the input to a network acts, unintentio nally, as the *affective* component supporting my assertion for this to be an intrinsic part of the learning model. This may seem odd to UK Psychologists however not to many computer scientists. “Affective computing is currently one of the most exciting areas of research in the cognitive and computing sciences….” (Johnson, 2001).

One such theory of affective computing (Picard, 1997) describes the importance of building an ‘emotion’ component into computers and argues that the computers’ emotions must emerge according to the computer’s own requirements (for what the computer will be used). Picard suggests such a level of representations that come together in any complete system for recognising, expressing or having emotions. She argues that emergent emotions in computing are attributable to systems based on their observable emotional behaviours and outward expressions and especially for *complex goal-directed autonomous systems*. That is motivational subroutines linking cognition to action where 'urges' are triggered in relevant situations. Picard argues that such activation will subsequently influence cognitive processes and attention. Here the connection between Damasio’s human emotion theory and Picard’s affective computing can be developed. Thereby, implementing the fast innate processing with the hard wiring of a computer, while the slower cognitively generated emotions would be implemented by the software:- fast but rigid and error prone followed by slow, mediated and flexibly more accurate activation. To continue with the Damasio analogy, the fast, rigid hardware must also be capable of ‘high-jacking’ the slow flexible processing. This idea was supported by Seidenberg *et al.* (1996) suggestion of two interacting learning systems. First is the fast, associative medial temporal system and the slow, integrative neocortical-type system. An early model of this type which incorporated emotion into the decision making process (Wright, Sloman & Beaudoin, 1995) using a three layer model to simulate the effect of grief on decision making processes. The layers were developed as an analogy of the brain, according to its phylogenic development and functional similarity. These layers comprise the reactive layer that executes automatically to the environment (e.g. new exemplars). It is fast (hard wired) and executes without prior conscious or cognitive appraisal (analogous to the limbic system and lower brain stem), followed by a deliberative layer (analogous to Damasio’s secondary processes) of cognitively generated emotions that typically require input from cortical reasoning, e.g. planning, evaluation, generalisation, etc. Finally, there is a self monitoring layer, analogous to many frontal lobe processes and is concerned with meta-management that prevent certain goals from interfering with each other to employ strategies and allocate resources. Thus, the idea of implementing an SSPS is not quite as far fetched as it may have first appeared.

The SSPS hypothesis was formed from observations from many sources and from the generic models of word, object and face processing. These explained steady state systems, and did not explain how a new exemplar would derive its own 'unit' or pattern of activation. For example, in logen-type model of word recognition, the grapheme to phoneme conversion route was incorporated to accommodate people's ability to read none words and to explain some single anomic effects. It does not distinguish between the ability to reproduce a none-word e.g. grertsey, via the grapheme to phoneme conversion route and the ability to learn a new foreign word, e.g. 'grertsey' (hello in Swiss). However, an individual would be able to set such into context and initiate specialised processing Swiss-language system.

Burton (1994) demonstrated that his backprop could learn new faces, however there was no rational for the *why* a certain new 'unit' should be learned. We do not need a new 'unit' for each new face encountered, furthermore, this model could not account for the robust AoA effects reported on face processing tasks. Fundamentally, SSPS proposes that new 'units' or patterns of activation are created by a combination of processes between external stimuli, the semantic system and the attention or arousal system, which raises *affect* level. For example, in one's environment, at any given time, there are a multitude of unattended stimuli, people become accustomed to 'ignoring' the sounds of traffic, music etc, however, as the general background noise changes one's attention is drawn to the change. I am arguing that one attends to new exemplars in an enhanced, but different way than to familiar stimuli. We monitor our environment via our physical senses and that the physical sensations feed one's semantic framework of the world. The semantic system is then more active in this process by actively monitoring one's perceptions of the environment, ready for changes. Such a change in the environment would automatically raise the level of *affect* as determined by the salience (attention grabbing) in the environment.

In this outline I argue that a rise in *affect* enhances learning new exemplars and sets up a specialised mechanism for processing the same type of information. The level of *affect* is determined by the salience creating valenced responses. From this two clear predictions arise: First, that an SSPS is created only when no pertinent information exists in the semantic system. Therefore, any AoA effects found on tasks exploring semantic memory be different to those previously reported. Second, that when a measure of *affect* (rated valence) is included as a variable, this measure will influence processing speeds in the same direction as AoA. These predictions cannot be derived from the other accounts AoA because they do not differentiate between the perceptual, semantic and output processing, or argue for an *affect* component.

Investigating AoA on semantic classification: Varied AoA effects are reported, however, little interest is paid to the null effects of AoA reported by Morrison *et al.* (1992) on a semantic task where participants classified line drawings for whether they depicted objects that were 'natural' or 'manmade'. Another paper often cited in support of the AoA effect (Rubin, 1980) is rarely correctly reported. For Rubin reported significant AoA effects, "… AoA is a good predictor of several semantic memory tasks, but not of recall or recognition" (pp 741). Contrary to other reports, Rubin reported AoA to significantly affect semantic classifications, but not word name production. However, the paper is very difficult to unpack as many separate studies are analysed together. Also, while naming speed and AoA were significantly correlated, Rubin argued that the regression analyses failed to support this. As with many early regression studies the intercorrelation statistics were not reported (or possibly investigated). The independent variables included seven measures of word frequency and one measure of AoA all of which were significantly intercorrelated (e.g. $r = .63$). Therefore, it is possible that measures of word frequency were confounded with AoA by introducing suppressor variables (see Moore & Valentine, 1998 for discussion).

Brysbaert, Van Wijnendaele and De Deyne (2000) report participants were significantly faster at generating semantic associates to early-acquired words than to late-acquired words. Their data was interpreted as demonstrating that AoA effects are located in the semantic system. However, this interpretation may have been premature, because the stimuli used were derived from a 1981 study, where teachers assessed what words six year olds should understand. The majority of AoA studies (especially Morrison and co-workers) use words first acquired from the age of three. Furthermore, Morrison *et al* (1997) validated the measures empirically, using child participants. It is also clear that language is flexible, and words that were rare or none-existent in 1981 are now very early-acquired, e.g. video, teletubbies, computer, etc, these controls not reported. Secondly, in Experiment 1, participants generated semantic associates to a word, however, the actual responses were not reported, therefore it is impossible to detect whether the associates generated were early or late-acquired words. For example, the word 'poes' (Dutch for cat) has many potential associates, e.g. dog, tiger, Australian mountain lion, etc. A convincing demonstration of an AoA effect would be if the generated associates were early-acquired words, also demonstrating control of phonemic power, word frequency, imageability, etc. Third, in Experiment 2, participants were required to classify printed words as belonging to a category of a given name (first name) versus none-first name. Participants were required to decide whether each word had a "definable meaning or was a first name". The response times present the problem here (early-acquired mean = 569ms vs late-acquired mean = 632ms) which are shorter than the response times for making lexical decisions to the same stimulus set (early faster by 25ms.; late faster by 14ms.), as previously reported by Brysbaert *et al* 2000. The rational for this experiment was that all the ‘word’ items belonged to the same category, which, it may be argued facilitated the faster response times. However, with the 144 stimuli were many other categories, e.g. concept words vs concrete words, animal vs plant, living vs non-living etc. Therefore, such an argument may be tenuous, if one follows the processing hierarchy logic that developed to account for processing speeds, then a lexical decision made to a word will be fast because only access to the recognition stage is required. Semantic classifications are slowly because an extra stage of processing is required. The speed of word naming is facilitated because two routes are employed, the normal route via semantics and concurrently the automatic activation of the grapheme phoneme conversion route. Therefore, to suggest that a task requiring access to semantics would take the same, and even less time than a lexical decision may be less plausible. A more parsimonious account of their data is that participants were looking whether the printed words were first names or not. Therefore, no conclusive empirical AoA effects on semantic tasks has been demonstration.

I have predicted that there should be an AoA effect on all perceptual and motor tasks. An SSPS will occur at an individual level based on a person’s prior knowledge. Therefore, AoA will not affect semantic tasks in the same way as perceptual and motor tasks. Moore (1998) employed a semantic classification task where participants decided whether a celebrity was famous for hosting a ‘chat-show’. A boarder-line significant result favoured *late-acquired* celebrities (mean early-acquired RT = 920ms; late-acquired RT = 875ms; $p<.06$). Printed names were not significant between early-acquired (mean RT = 851ms) and late-acquired items (mean RT = 861ms; $p=.35$). Therefore, the robust advantage for early-acquired celebrities in face naming; FDT to faces or names and reading the names aloud was not apparent in the semantic classification task. However, this may have been due to the complication of the task (e.g. deciding whether celebrities were famous for appearing as chat-show-hosts[4]).

Unless otherwise stated, the experiments reported below shared the following features. Aged between 18 and 25 years 24 students participated in each study. Each took part in one experiment only. Testing was by computer, using the MEL2 experiment generator (Rodgers, Schneider, Pitcher & Zuccolotto, 1995) to program, randomise stimulus presentation, record RTs and accuracy of response.

The critical items were 50 pictures of celebrity’ faces or their printed names. Digitally scanned pictures created high contrast grey scale images (256 x 256 pixels). Printed names appeared in point 18 Arial font.

There were 25 early-acquired celebrities and 25 late-acquired celebrities. The two groups of items (early vs. late) were matched on other influential variables, i.e. the number of phonemes in the name; the phonemic power of the initial phoneme; rated distinctiveness of face and rated familiarity (frequency of encounter) with the celebrity. We created the two groups so that they differed significantly ($p<.01$) in the AoA ratings, but not on the other variables. The filler items created a set where the 50% would fulfil the criteria required by the experimental task.

All of the experiments use a within-participant factorial design with two levels (early vs late AoA). The DV was response latencies. Accuracy was monitored to ensure no speed/accuracy trade off occurred. Participants focused on the centre of the screen and decided as quickly and accurately as possible whether the stimuli fulfilled the experimental criteria, e.g. in Experiment 1, an occupation category appeared on the screen first, followed by either a celebrity face (condition 1) or their printed name (condition 2). Participants decided whether this was correct or not, and press the appropriate button 'Yes' or 'No'. During twelve practice trials participants were encouraged to respond as quickly as possible. To test the validity of stimulus group selection, participants rated the stimuli for familiarity, distinctiveness and AoA (these rating procedures are reported in Moore & Valentine, 1998). Analyses of the *post hoc* ratings are reported only where the factorial group distinction was violated.

Experiment 1 overcame the complicated task of "chat-show-host" because here an occupation-type was read before stimulus presentation so processing was 'on-line' as soon as the stimulus was recognised. There was no significance between the RTs to the faces of early-acquired (mean = 828ms.) and late-acquired celebrities (mean = 863ms.) or the names of early-acquired (mean = 856ms.) and late-acquired celebrities (mean = 823ms.). These results do not support a semantic affect on face or name processing, although these stimuli evinced a significant advantage for processing early-acquired faces and names on naming and FDT tasks to the same stimuli, with the same population.

Experiment 2 explored whether a semantic relationship between celebrities where the face (condition 1) or name (condition 2) of a celebrity double act was presented on the screen, followed by the face (or name) of the partner. Participants responded 'yes' to the second stimuli if it made up a pair with the first celebrity, otherwise the answer was 'no'. E.g., see Eric Morcombe respond to Ernie Wise (a famous contemporary UK double act). There was no significance between RTs of early-acquired (mean = 805ms.) and late-acquired celebrity' faces (mean = 829ms.) or to early-acquired (mean = 814ms) and late-acquired names (mean = 819ms.). A significant difference in *post hoc* familiarity ratings of early and late-acquired celebrities showed that late-acquired celebrities were significantly more familiar than early-acquired celebrities, this may account for the small advantage on early-acquired RTs.

Experiment 3 exploited the collected *post hoc* ratings to create new sets of critical items with a wider spread of AoA scores and used only the names of celebrities. The task was the same as Experiment 2, to decide whether two names of celebrities formed a pair. There was a highly significant difference between the classification latencies to the names of early-acquired (mean = 638ms) and late-acquired celebrities (mean = 577ms.; p<.001; and items p<.04). However, the advantage was for *late-acquired* names.

Some names employed here may have had an order-effect, e.g. 'Eric Morcombe' followed by 'Ernie Wise' is the usual presentation of this act, 'Morcombe and Wise'. This was examined by reversing the order of presentation (e.g. in condition 2 the first presentation was 'Ernie Wise' and the response made to 'Eric Morcombe'). There was no significant effect between early-acquired (mean = 616ms) and late-acquired celebrities (mean = 623ms.). Using identical items and task the results were not replicable when using unusual name pairs. These tasks required an explicit semantic classification to be made, i.e. is this celebrity in someway connected to one that had gone before. In a final attempt to capture AoA effects on a semantic task celebrities pairs were matched, as before, however, in **Experiment 4** the task was made implicit. Participants read the first name, but made no response, instead they responded only on the presentation of the second name.

The participants decided whether the second name was famous or not in a FDT. There was no effect between early-acquired (mean = 671ms) and late-acquired names (mean = 683ms.). The *post hoc* familiarity ratings revealed a significant advantage for early-acquired stimuli (i.e. more familiar than late-acquired), which could have accounted for the none-significant RT advantage. Table 1, shows the results of these experiments. The only significant AoA effect on semantic tasks was for *late-acquired* stimuli. These findings support the non-semantic, multiple perceptual and output loci of the SSPS hypothesis.

Table 1: Showing the mean reaction time and significance level for various semantic tasks to Celebrity' faces and names

Classification	Stimuli	Age of acquisition		
		Early	Late	Sig.
Previously reported data	Face	920	**875**	*p*<.06
'Chat-show-host'	Name	**851**	861	ns
Exp 1: Occupation	Face	**828**	863	ns
	Name	856	**823**	ns
Exp 2: Form a pair?	Faces	**805**	829	ns
	Names	**814**	819	ns
Exp 3: Form a pair?	New stimuli.	638	**577**	*p*<.001
	Reversed names	**616**	623	ns
Exp 4: FDT	Names	**671**	683	ns

The SSPS hypothesis alone would predict an effect of '*affect*' on naming celebrity' faces. There is some evidence in the literature to suggest that an effect of *affect* may be demonstrated by face recognition impairments, i.e. prosopagnosia. Once again, the allusion to multiple cerebral pathways connecting the visual cortex to the temporal lobe and subcortical limbic system/s. Bauer (1984: Bauer & Verfaellie, 1988) argued for two functionally and anatomically separate (ventral and dorsal) routes mediating face recognition. The ventral route subsumes the function of overt recognition because lesions to this pathway would disconnect facial perception from associated memories. The dorsal route subsumes the function of recognition by activating rapid, selective orientation to the face thus inducing emotional arousal. Bauer proposed this route to be intact in prosopagnosia and accounts for the significant raise in skin conductance responses (SCRs) to familiar faces, even when the faces were presented covertly (e.g. Tranel & Damasio, 1985). What follows are some preliminary explorations into the effect of *affect* on face processing. A correlational design was employed following Moore and Valentine (1998) initial AoA explorations.

Experiment 5 Here 20 participants rated 100 celebrities on familiarity (as above). The 72 with the highest familiar ratings (5+) were selected. Twenty-four participants named these faces followed by rating the faces as before, and for a measure of *affect*.

Affect ratings used a 14 point scale:- -7 -6 -5 -4 -3 -2 -1 0 +1 +2 +3 +4 +5 +6 +7

Here –7 indicates a celebrity for whom a strong negative feeling is attributed: 0 indicates a celebrity for whom no feeling is attributed: 7+ indicates a celebrity for whom a strong positive feeling is attributed. For analyses, positive and negative signs were removed and ratings collapsed to produce a 7 point scale reflecting the magnitude of *affect*, i.e. 0 for a celebrity for whom no *affect* is attributed to 7 for strong attribution of *affect*.

Thirty seven per cent of responses were removed as invalid, 20 items were removed because of high error score (over 50%) leaving 52 faces. Reciprocal transformations reduced a long skew (typical of face naming) in the RTs reported as naming speed. The descriptive data and correlation statistics appear in Table 2 to reveal all of the IV's to be highly correlated with naming speed, most were also intercorrelated showing that highly-familiar celebrities were also rated as being more distinctive than less-familiar celebrities. Highly-familiar celebrities were rated as creating a higher *affect* response than less-familiar celebrities. A negative correlation revealed that early-acquired celebrities were more distinctive than were late-acquired celebrities. AoA did not significantly intercorrelate with familiarity (as reported in Moore & Valentine, 1998) nor with the measure of *affect* (the lowest intercorrelation of all). With such high intercorrelations, interpretation must be treated with caution[8]. In the simultaneous regression 46% of the variance was accounted for by all of the IV's R^2 =.47; S.E. = 154.68; $F(4,47) = 10.59$; MSE = 23950; $p < .0001$. The three standardised regression coefficients with naming speed were *affect* ($p<.001$), distinctiveness ($p<.05$) and a borderline AoA ($p<.067$).

Table 2: Showing the mean naming latencies, mean rating scores raw correlations and multiple regression.

Variables	**1 N. Spd**	**2 Fam.**	**3 Dist.**	**4 AoA**	**5 Affect**	***B***	**β**
1. (DV) N.Sp		.50**	-.46**	.31**	-.55**		
2. Familiarity			.30*	-.21	-.59**		
3. Distinctiveness				-.36**	.24*	-58.456	-.270*
4. Age of Acquisition					.03	66.886	.213
5. Affect						-138.191	-.491**
						Intercept =	08
Means	1474 RT	4.58	3.52	4.42	3.34		
	205	0.45	0.95	0.65	0.73		
							R^2 = .47
						Adjusted	R^2 = .43
							R = .68**
							S.E. = 155

N.Sp. = Naming Speed *= p<.05 **= p<.01

B = Unstandardize Regression Coefficients; β = Standardised Regression Coefficients

[8] As the IV's are intercorrelated 0.15 to enter and 0.20 to remove was selected, as recommended by Bendel and Afifi (1977) so that otherwise important IV's are less likely to be excluded.

In the stepwise regression, *affect* entered first, accounting for 31% of the variance in naming speed (R^2 =.31; SE = 172.22; $F(1,50) = 22.17$; MSE = 29659.96; $p<.0001$). Second to enter was distinctiveness (R^2[change] =.12; SE = 158.57; $F(2,49) = 18.07$; MSE = 25143.53; $p<.0001$). Finally, AoA entered (R^2[change] =.04; SE = 154.68; $F(3,48) = 13.82$; MSE = 23925.26; $p<.0001$). Therefore, variance in naming speed was accounted for by *Affect* (31%), Distinctiveness (12%) and AoA (4%). However, the intercorrelation statistics were valid for only *Affect* and Distinctiveness (Tolerance [Tol.]= .944; Variance inflation factor [VIF] = 1.06). When AoA was entered, the model was less robust (Tol = .859; VIF = 1.164).

Finding no Familiarity effect may be because of the manipulation before the experiment (using highly familiar items), to test this a further stepwise regression was run omitting the *affect* variable. Familiarity entered first accounting for 25% of the variance (R^2 =.25; SE = 179.70; $F(1,50) = 16.29$; MSE = 32291.17; $p<.0001$). For brevity the accuracy data are not presented, however, the same pattern of results also occurred.

I have discussed the problems of interpreting multiple regression data elsewhere. For now, a factorial analysis of the data may help to clarify these results. The data were divided based on a median split of *affect* measures. The groups were divided in such a way as to match on the other variables and differ on measures of *affect*. The group statistics appear in Table 3, and demonstrates that *affect* may be an importance measure to consider in face processing studies.

Table 3: High and low *affect* group statistics.

	High Affect	Low Affect	Difference
Familiarity	4.80 (.42)	4.37 (.37)	ns
Distinctiveness	3.73 (.92)	3.31 (.94)	ns
Age of acquisition	4.45 (.75)	4.38 (.55)	ns
Affect	3.95 (.44)	2.43 (.35)	$F(1,50)=11.09$; $p<.005$
Naming Latencies	1379 (182)	1570 (183)	$F(1,50)=14.19$; $p<.001$
Accurate Responses	20.08 (2.67)	17.58 (2.96)	$F(1,50)=10.25$; $p<.005$

Experiment 6: The effect of *affect* above was found on an explicit processing task, however it is interesting that prospagnosics preserved ability is demonstrated on implicit not explicit face recognition (e.g. patient LF, De Hann, Bauer & Greve, 1992; also see Tranel, Damasio, & Damasio (1995). Stone, Valentine and Davies (2001) argued that effects of implicit face processing, similar to those demonstrated

for prosopagnosia should be present in the intact brain. Using SCR as a measure of recognition Stone *et al* tested recognition between familiar and unfamiliar celebrities at conscious (supra-threshold, i.e. 220ms) and sub-conscious (sub-threshold, i.e. 16ms) levels. Based on rated valence, Stone et al used famous faces of 'good' (e.g. Nelson Mandela), 'neutral' (e.g. J.F. Kennedy), and 'evil' (e.g. Adolf Hitler). Participants viewed faces at sub- and supra-threshold exposures and generated the names followed by *post hoc* ratings of familiarity, AoA, valence and distinctiveness. They report AoA uniquely predicted 38% of the variance, *affect* 18% and Familiarity 6.5%. The intercorrelation between familiarity, AoA and distinctiveness was not significant, however, they did contribute to the shared and unique variances accounting for 63% of the overall variance in the data. We investigated the degree to which *affect* could facilitate naming faces presented at optimal covert thresholds.

Fifty participants rated 250 celebrity' names and 15 common celebrity occupations for familiarity. Thirty celebrities and six common occupations with the highest scores were selected. Eighty students were pseudo-randomly allocated to participate in each condition.

There were 30 critical faces, and 30 filler faces, all chosen according to common occupation. There were 5 sets of common occupations (e.g. sport, film, etc.) with 6 faces in each in a between-participants design to compare two conditions with *targets present* (Condition I) vs *targets absent* (Condition II). The DV was the accuracy of response.

The time of stimuli exposure was calibrated individually in Condition I with exposure durations (16ms, 30ms, 50ms, 66ms and 80ms respectively) beginning at 16ms and increased until participants were able to report seeing a line drawing. Once conscious exposure level was established, exposure duration was set to the previous level Condition II used exposure of 16ms only.

Participants focused on the centre of the screen where 5 flashes were followed by a mask and a question appeared, requiring the generation of "any famous names that come to mind (with a maximum of 6)" or "any celebrity occupation" with 5 question of each. Participants in Condition II had 16ms exposure duration without calibration. They were informed that famous faces would be flashed subliminally. No stimuli were presented, only the flashes, mask and questions. All participants provided *post hoc* ratings.

Participants' responses were scored as i.) correct names; ii.) [given in the] correct place; iii.) correct occupation; and iv.) [given in the] correct place.

Table 4: Response statistics.

Answer Type	I Targets Present II Targets Absent	Mean	S.D.	t	df	Sig. *p*<.
	Present	11.00	4.96	-3.617	65	.001
	Absent	16.40	6.60			
correct name, correct place	Present	2.00	2.39	3.24	78	.002
	Absent	.68	.997			
correct " , wrong place	Present	1.18	1.22	3.06	78	.001
	Absent	2.55	2.09			
correct occupation, correct place	Present	1.85	1.29	5.43	78	.0001
	Absent	.60	.67			
correct ", wrong place	Present	1.10	.93	4.96	78	.0001
	Absent	2.28	1.18			

The number of responses in Target Absent condition were significantly greater than those provided by participants who were presented with faces. However, these were in the wrong place significantly more often than participants who were exposed to the celebrities. Participant in Condition I produced significantly more correct names (and occupations) at the correct place than participants who had not been exposed to the faces. A regression analysis was performed however, a computer error deleted distinctiveness rating data therefore the following analyses are merely an *indication* of the potential of *affect* as a predictor variable. The number of correct responses in the Target Present condition were entered into a regression analysis, together with rating scores of familiarity (mean scores = 5.23), AoA (mean score = 5.38) and *affect* (mean score = 3.10). All variables significantly correlated (familiarity, $p<.001$; AoA, $p<.001$; *affect*, $p<.001$) with the DV to demonstrate that more correct responses occurred to celebrities rated as highly-familiar, early-acquired and with strong *affect* than occurred for less-familiar celebrities, later-acquired with lower ratings of *affect*.

The significant intercorrelations between ratings of familiarity and AoA ($p<.001$) and *affect* ($p<.001$) showed that celebrities rated as highly familiar were acquired early in life and generated a higher *affect* score than less familiar faces. The simultaneous regression suggested that 26% of the variance was accounted for by those variables (R^2 =.26; SE = 1.90; $F(1,59) = 16.82$; $p<.001$). However, the collinearity statistics were incompatible with accepting this model (Tol .289, VIF = 3.46). These data suggest that participants who were exposed to faces produced more of the correct names and occupations in the correct places, than did participants who were guessing. The significant correlations and intercorrelations clearly indicate the potential of *affect* such studies.

The data retorted here refute a *single* locus for AoA effects at the phonological output lexicon or in the semantic system. Instead, these data are consistent with the view that AoA effects are the result of the neural representation of any class of perceptual or lexical information. This, in turn, may reflect a general property of the mental representation of perceptual information both visual and lexical. Moore (1998; 2000) argued that *all* new patterns of information are processed in a fundamentally different way to later-acquired related material because initial encounters of exemplars from *any* new class of information, of *any* type and *at any age*, could begin to establish a new principle for organising and processing the relevant information. If enough exemplars of a class are encountered, then a specialised processing system would become established. Later-acquired, related information would be represented in a different manner from early items because the early-exemplars were involved in setting out the parameters for a dedicated processing mechanism. Under this view the 'phonological mapping' hypothesis may be *one* such dedicated system. However, this effect arises because of the active involvement where initial exemplars '*configure'* the system. There is some evidence for such *facilitation* in the Ellis and Lambon-Ralph model, as reported by Monaghan and Ellis (2002). Monaghan and Ellis demonstrated that late-learned patterns could take advantage of the network's configuration, but this was only when a high degree of correlation existed between early and late-learned patterns. The evidence for the distinction between network configuration and reduced plasticity may be seen by the fact that while later-learned patterns are often more difficult, they *are* indeed learned. Furthermore, people become proficient, even expert, in specific fields where the information has been learned later in life, as demonstrated in the way specialist expertise can be acquired. For example, taxi drivers superior navigational skills and how their knowledge correlated with neural structures in the posterior hippocampi (Maguire *et al*, 2000). This proficiency was also demonstrated in the special expertise of learning a second language (Pernani, Paulesu, Galles, Dupoux, Dehaene, Valentino, Cappa, Fazio & Mehler, 1998; Flege, Yeni-Komshiam & Lui, 1999; Birdsong & Molis, 2001). The challenge presented to neural networks researchers is the question of how the network comes to be established in the first place, and to account for developing expertise after substantial learning has already taken place, as in the case of navigational skill and second language learning.

Conclusion: The ideas presented here are not new. The novelty is in bringing together disparate areas of research in an attempt to create a unified theory of *how* and *why* certain empirical effects on words, objects and people sit comfortably together. This is in line with many current attempts to explicitly introduce the concept of the mind, emotion/affect, into general cognitive processing.

Wide areas of research have provided convergent support. I hope that it is clear that the SSPS hypothesis incorporates models of frequency, the phonological completeness and neural plasticity. It is no surprise that neuropsychological data reveal amnesics to evince preserved islands of memory for information that relates to their previous personal interests (e.g. Verfaellie *et al*, 1995; Baddeley, 2001; Kitchener & Hodges, 1999). If the SSPS hypothesis is valid, it may be a useful rehabilitative tool for other cognitive impairments.

This outline of the SSPS argues that a rise in affect will influence the learning new exemplars of information and set up a specialised mechanism for processing future exemplars of the same type of information. The level of *affect* is determined by the salience of exemplars within the environment. In this I go further than explaining AoA effects instead the focus is an attempt to create a unified account for a variety of phenomena. Therefore, several further assumptions are incorporated. i.) Individuals begin their oriented learning in utero, especially auditory information. ii.) Infants are born with certain traits of their individual personality that may have been developed by sensory experiences in-utero. iii.) During this time, infants habituate to stimuli that are repeatedly in their immediate environment. Following a period habituation, infants orientate to new information, as demonstrated by the 'sucking or looking' at new stimuli presented. iv.). Temporal frequency is important so that as more information of the same kind is mapped onto previous knowledge, increasing neural connectivity as demonstrated by animals in reared in enriched environments developing grater cortical mass that those in a sterile environment (Bliss, 1999). v.) That, learning of sufficient exemplars of 'new information' will physically peg-out the parameters for that ilk of information, therefore, early-learned material will be processed explicitly, whereas later-learned material will be incorporated implicitly.

References.

Anderson, A. K. & Phelps, E. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*. 411, 306-309.

Baayen, H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Philadelphia: PA: University of Pennsylvania Linguistic Data Consortium.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*.

Bauer, R. M. (1984). Automatic recognition of names and faces in prosopagnosia: A neuropsychological application of the guilty knowledge test. *Neuropsychologia*. 22, 457 – 469.

Bauer, R. M. & Verfaellie, M. (1988). Electrodermal discrimination of familiar but not unfamiliar faces in prosopagnosia. *Brain and Cognition*. 8, 240-252.

Bendel, R.B. & Afifi, A.A. (1977). Comparison of stopping rules in forward multiple regression. *Journal of American Statistical Association*. 72, 46-53.

Birdsong, D. & Molis, M. (2001). On evidence for maturational constraints in second language acquisition. *Journal of Memory and Language*. 44, 235-249.

Bliss, T.V.P. (1999). Young receptors make smart mice. *Nature*. 401, 25-27.

Bonin, P., Chalard, M., Méot, A. & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*. 93, 89-114.

Bonin, P., Fayol, M. & Chalard, M. (2001). Age of acquisition and word frequency in written picture naming. *Quarterly Journal of Experimental Psychology*. 54, 469-489.

Bradley, M. M., Cuthbert,B. N. & Lang, P. J.(1993). Pictures as prepulse: Attention and emotion in startle modification. *Psychophysiology*. 30: 541-545.

Brown, G.D.A., & Watson, F.L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory and Cognition*. 15, 208-216.

Brysbaert, M., Lange, M. & Van Wijnendaele, I. (2000). The AoA effects and frequency-of-occurrence in visual word recognition: Further evidence for Dutch. *European Journal of Cognitive Psychology*. 12, 65-86.

Brysbaert, M., Van Wijnendaele, I. & de Deyne (2000). Age of acquisition effects in semantic processing tasks. *Acta Psychologica*. 104, 215-226.

Burton, A. M. (1994). Learning new faces in an interactive activation and competition model. *Visual Cognition*. 1, 313-348.

Burton, A. M., Bruce, V. & Johnson, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*. 81, 361-380.

Bruce, V. & Valentine, T. (1985). Identity priming in the recognition of familiar faces. *British Journal of Psychology*. 76, 373-383.

Bruce, V. & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology*. 77, 305-327.

Carroll, J. B. & White, M. N. (1973a). Word frequency and AoA as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*. 25, 85-95.

Cuetos, F., Ellis, A. W. & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments and Computers,* 31. 650-658.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Proceedings of the Royal Society of London*. 351, 1413-1420.

Damasio, A. R. (1999). *The Feeling of What Happens: Body Emotions and Making of Consciousness*. New York: Harcourt Brace.

DeCasper, A.J. & Spence, M.J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development,* 9, 133-150

Ellis, A.W. & Lambon Ralph, M.A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory and Cognition.* 26, 1103-1123.

Ellis, A.W. & Morrison, C.M. (1998). Real AoA effects in lexical retrieval. *Journal of Experimental Psychology: Learning Memory and Cognition.* 24, 515-513.

De Haan, E. H. F., Bauer, R. M., & Breve, K. W. (1992). Behavioural and physiological evidence for covert face recognition in a prosopagnosic patient. *Cortex*. 28, 77-95.

Flege, J.E., Yeni-Komshiam, G.H. & Lui, S. (1999) Age constraints on second language acquisition. *Journal of Memory and Language.* 41, 78-104.

Graham, F. K. (1980). Control of reflex blink excitability. In F. Thompson, L. Hicks & V. B. Shvyrkov (Eds), *Neural Nechanisms of Goal Directed Behaviour and Learning.* 511-519. Academic Press. New York.

Gerhand, S. & Barry, C. (1999). Age of acquisition, frequency and the role of retrieved phonology in the lexical decision task. *Memory & Cognition.* 27, 592-602.

Gilhooly, K.J. (1984). Word AoA and residence time in lexical memory as factors in word naming. *Current Psychological Research and Reviews.* 3, 24-31

Gilhooly, K.J. & Gilhooly, M.L.M. (1979). Age of acquisition effects in lexical and episodic memory tasks. *Memory and Cognition.* 7, 214-223.

Goldfield, B. & Reznick, J. S. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology*. 28, 406-413.

Hepper P G. (1991). An examination of fetal learning before and after birth. *Irish Journal of Psychology*; 12, 95-107.

Izzard, C. E. (1993). Four systems of emotion activation: Cognitive and noncognitive processes. *Psychological Review.* 100: 68-90.

Jescheniak, J. D. & Levelt, W.J.M. (1994). Word frequency in speech production: Retrieval of syntactic information and phonological form. *Journal of Experimental Psychology: Learning Memory and Cognition*, 20, 824-843.

Johnson, C. (2001). The introduction to the AISB'01 Symposium on Emotion, Cognition and Affective Computing. *The Society for the study of Artificial Intelligence and Simulation of Behaviour.* ISBN 11 902956 19 7. pp iii.

Juscyk, P. W. & Aslin, R. N. (1995). Infant's detection of sound patterns of words in fluent speech. *Cognitive Psychology*. 29, 1-23.

Juscyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development.* 64, 675- 687.

Killcross, S (1999). The amygdala, emotion and learning. The Spearman Medal Lecture. *The Psychologist* (2000); 13, 502-507.

Kitchener, E.G. & Hodges, J.R. (1999). Impaired knowledge of famous people and events with intact autobiographical memory in a case of progressive right temporal lobe degeneration: Implications for the organisation of remote memory. *Cognitive Neuropsychology*. 16, 589-607.

Kokelaar, S. (2000). The rhythem of language. *Trends in Cognitive Sciences.* 4, 6: 214.

Kuèera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, 50**,** 372-385.

Lang, P.J., Bradley, M.M. & Cuthbert, B.N. (1997). Motivated attention: Affect, activation and action. In P. J. Lang, R. F. Simons & M. T. Balaban (Eds.). *Attention and Orienting: Sensory and Motivational Processes*. 97-136. Hillisdale, NJ: Erlbaum.

Langer, E. J. (200). Mindful Learning. *Current Directions in Psychological Science. 9*, (6). 220-223.

Le Grand, R., Mondloch, C.J., Maurer, D. & Brent, H.P. (2001). Early visual experiences and face processing. *Brief Communications; Nature*, 410, 890.

Le Doux, J. E. (1990). Information flow from sensation to emotion: Plasticity in the neural computation of stimulus value. In M. Gabriel and J. Moore, (eds) *Learning and Computational Neuroscience: Foundations of Adaptive Networks,* MIT Press.

Le Doux, J. E. (1994). Emotional memory and the brain. *Scientific American,* June 1994, 50-57.

Le Doux, J. E. (1996). *The Emotional Brain.* Simon & Schuster, New York.

Lewis, M.B. (1999). Age of acquisition in face categorisation: Is there an instance based model account? *Cognition*, 71, 23-39.

Maguire, E.A., Gadian, D.G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S.J. & Frith, C.D. (2000). Navigation-related structural change in the hippocampi of taxis drivers. *Proceedings of National Academic of Sciences (USA),* 10, 1073 – 1077.

Mayes A. R., Downes, J. J., McDonald, C., Poole, [Moore] V. M., Rooke, S., Sagar, H. J., & Meudell, P.R. (1994). Two tests for assessing remote public knowledge: A tool for assessing retrograde amnesia, *Memory.* 2, 183-210.

Mayes A. R., Poole, [Moore] V. M. & Gooding, T. (1991). Increased reading speed for words and pronounceable non-words: Evidence of preserved priming in amnesics. *Cortex*. 27, 403-415.

Monaghan, J. & Ellis, A.W. (2002). What exactly interacts with spelling–sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory and Cognition.* 28, 183–206.

Monsell, S., Doyle, M.C. & Haggard, P.N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General.* 118, 43-71.

Moore, V. (1998). *The effects of age of acquisition in processing people's faces and names.* Thesis submitted for the Degree of Doctor of Philosophy, University of Durham.

Moore, V. (2000). *Convergent evidence for the mechanism proposed to account for the effects of AoA.* Paper presented: XXVII International Congress of Psychology (Stockholm).

Moore, V. & Valentine, T. (1998). Naming faces: The effect of AoA on speed and accuracy of naming famous faces. *Quarterly Journal of Psychology, 51,* 485-513.

Moore, V. & Valentine, T. (1999). The effects of AoA in processing famous faces and names: Exploring the locus and proposing a mechanism *Proceedings of the XXI Annual Meeting of the Cognitive Science Society, Vancouver*, 1999, 416-421. Mahwah: Lawrence Erlbaum Associate.

Moore, V. & Valentine, T. (2001). *Convergent evidence to support the Set up of a Specialised Processing System hypothesis*. Paper presented in symposium on Age of Acquisition for the joint meeting of the British Psychological Society Cognitive Section and the European Society for Cognitive Psychology conference in Edinburgh 5-9th Sept.

Moore, V., Smith-Spark, J. Valentine, T. (in press). The effects of AoA on perceptual object recognition. *European Journal of Cognitive Psychology.*

Moore, V., Valentine. T. & Turner, J. (1999). Age-of-acquisition and cumulative frequency have independent effects. *Cognition.* 72, 305-309.

Morrison, C.M., Chapell, T. & Ellis, A.W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology.* 50, 528-559.

Morrison, C.M. & Ellis, A.W. (1995). The roles of word frequency and AoA in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory & Cognition.* 21, 116-133.

Morrison, C.M., Ellis, A.W. & Quinlan, P.T. (1992). Age of acquisition, not word frequency, affect object naming not recognition, *Memory and Cognition.* 20, 705-714.

Morrison, C. M., Hirsh, K.W. Chapell, T. D. & Ellis, A.W. (2002). Age and AoA: An evaluation of the cumulative frequency hypothesis. *European Journal of Cognitive Psychology.* 14, 435-459.

Oldfield, R.C. & Wingfield, A. (1964). The time it takes to name an object. *Nature.* 202, 1031-1032.

Oldfield, R.C. & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*. 17, 273-281.

Ohman, A. (1996). Preferential pre-attentive processing in anxiety: Preparedness and attentional biases. In R. M. Rapee (Ed.), *Current Controversies in the Anxiety Disorders.* 253-290. New York: Guilford.

Ohman, A. & Soares, J. F. (1994). "Unconscious anxiety": Phobic responses to masked stimuli. *Journal of Abnormal Psychology*. 103: 231-240.

O'Reilly, R.C. & Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding theMmind by Simulating the Brain.* MIT Press, Cambridge Massachussetts.

Pernani, D., Paulesu, E., Galles, N.S., Dupoux, E., Dehaene, S., Valentina, V., Cappa, S.F., Fazio, F. & Mehler, J. (1998). The bilingual brain: Proficiency and AoA of the second language. *Brain*. 121, 1841-1852.

Picard, R. (1997). *Affective Computing*. MIT Press. Cambridge, Massachusetts, London, England.

Rinsland. H.D. (1945). *A Basic Vocabulary Of Elementary School Children.* Macmillan: New York.

Rodgers, K.A., Schneider, W., Pitcher, E. & Zuccolotto, A. (1995). Mel Professional. Pittsburgh, PA: Psychology Software Tools, Inc.

Rubin, D. C. (1980). 51 properties of 125 words: A unit analysis of verbal behaviour. *Journal of Verbal Learning and Verbal Behaviour.* 19, 736-755.

Scarborough, D.L., Cortese, C. & Scarborough, H.S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology.* 3, 1-17.

Seidenberg, M.S. & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*. 96, 523-68.

Seidenberg, M.S., Peterson, A., MacDonald, M.C. & Plaut, D.C. (1996). Pseudohomophones effects and models of word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition.* 22, 48-62.

Sergent, J. & Poncet, M. (1990). From Covert to overt recognition in a prosopagnosic patient. Brain. 113, 989-1004.

Shallice, T. & Kartsounis, L.D. (1993). Sective impairment of retrieving people's names: A category specific disorder? *Cortex*. 93, 281-291.

Simon H. A. (1967). Motivational and emotional controls of cognition. In *Models of Thought*. Yale University Press, New Haven, 1979.

Stone, A., Valentine, T. & Davies, R. (2001). Face recognition and emotional valence: Processing without awareness by neurologically-intact participants does not simulate covert recognition in prosopagnosia. *Cognitive, Affective and Behavioural Neuroscience.* 1, 183-191.

Tarr, M.J. & Cheng, Y.D. (2003). Learning to see faces and objects. *Trends in Cognitive Sciences.* 17, 23-30.

Thorndike, E.L. & Lorge, L.I. (1944). *The Teacher's Wordbook of 30,000 Words*. New York: Teachers College Press.

Tipples, J. & Sharman, D. (2000). Orienting and affective pictures. *British Journal of Psychology*. 91 (1), 85-98.

Toda, M. (1982). Design of a fungus-eater. *Behaviour Science.* 7, 164-183.

Tranel, D. & Damasio, A.R. (1985). Knowledge without awareness: An automatic index of facial expression by prosopagnosics. *Science*. 228, 1453-1454.

Tranel, D. & Damasio, H.& Damasio A.R. (1995). Double dissociation between overt and covert face recognition. *Journal of Cognitive Neuroscience.* 7, 425-432.

Turner, J.E., Valentine, T. & Ellis, A.W. (1998). Contrasting effects of Age of acquisition and word frequency on auditory and visual lexical decision. *Memory and Cognition*. 26, 1282-1291.

Valentine, T., Brennen, T. & Brédart, S. (1996). *The Cognitive Psychology Of Proper Names. On The Importance Of Being Ernest*. London: Routledge.

Valentine, T., Hollis, J. & Moore, V. (1998). On the relationship between reading, listening and speaking: It's different for people's names. *Memory and Cognition.* 26, 740-753.

Valentine, T., Moore, V., Flude, B., Young, A. W. & Ellis, A. E. (1993). Repetition Priming of proper name processing. Do common nouns and proper names prime each other? *Memory.* 1, 329-349.

Verfaellie, M., Croce, P., & Milberg, W. P. (1995). The role of episodic memory in semantic learning: An examination of vocabulary acquisition in a patient with amnesia due to encephalitis. *Neurocase*, 1, 291-304.

Werker, J. F. (1994). Cross language speech perception: developmental change does not involve loss. *In The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words.* 167-224. J. C. Godman and H. C. Nusbaum (Eds.), MIT Press.

Woodward, A. L., Markman, E. M. & Fitzsimmons, C. M. (1994). Rapid word learning in 13- & 18-month-olds. *Developmental Psychology.* 30, 553-566.

Wright, I., Sloman, A. & Beaudoin, L. (1996). Towards a design-based analysis of emotional episodes. *Philosophy, Psychiatry and Psychology.* 3, 101-126.

Yamazaki, M. Ellis, A. W., Morrison, C. M., & Lambon Ralph, M. A. (1997). Two age of acquisition effects in the reading of Japanese Kanji. *British Journal of Psychology.* 88, 407-421.

Yantis, S. (1995). Attentional capture in vision. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds), *Converging operations in the study of visual selective attention.* 45-77. Washington, DC: American Psychologist Association.

Zevin, J.D. & Seidenberg, M.S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language.* 47, 1-29.