



Frederico Miranda Rodrigues Pinheiro

**O Papel da Informação Semântica na Atenção Seletiva:
O uso de dicas endógenas como textura em uma tarefa de
busca visual**

**UBERLÂNDIA
2011**



Frederico Miranda Rodrigues Pinheiro

**O Papel da Informação Semântica na Atenção Seletiva:
O uso de dicas endógenas como textura em uma tarefa de
busca visual**

Dissertação apresentada ao Programa de Pós-Graduação em Psicologia – Mestrado, do Instituto de Psicologia da Universidade Federal de Uberlândia, como requisito parcial à obtenção do Título de Mestre em Psicologia Aplicada.

Área de Concentração: Psicologia Aplicada

Orientador: Dr. Joaquim Carlos Rossini

**UBERLÂNDIA
2011**



Frederico Miranda Rodrigues Pinheiro

O Papel da Informação Semântica na Atenção Seletiva: O uso de dicas endógenas como textura em uma tarefa de busca visual

Dissertação apresentada ao Programa de Pós-Graduação em Psicologia – Mestrado, do Instituto de Psicologia da Universidade Federal de Uberlândia, como requisito parcial à obtenção do Título de Mestre em Psicologia Aplicada.

Área de Concentração: Psicologia Aplicada

Orientador: Dr. Joaquim Carlos Rossini

Composição da Banca

Dr. Prof. Joaquim Carlos Rossini (Orientador)

Dr. Prof. Ederaldo José Lopez (Docente da UFU)

Dr. Prof. César Alexis Galera (Docente da Universidade de São Paulo)

Dra. Susi Lippi Marques Oliveira (Docente da Universidade de São Carlos)

**UBERLÂNDIA
2011**

Dados Internacionais de Catalogação na Publicação (CIP)

Sistema de Bibliotecas da UFU, MG, Brasil.

P654p Pinheiro, Frederico Miranda Rodrigues, 1985-
 O papel da informação semântica na atenção seletiva [manuscrito] : o
 uso de dicas endógenas como textura em uma tarefa de busca visual /
 Frederico Miranda Rodrigues Pinheiro. – 2011.
 83 f.

 Orientador: Joaquim Carlos Rossini.
 Dissertação (mestrado) - Universidade Federal de Uberlândia,
 Programa de Pós-Graduação em Psicologia.
 Inclui bibliografia.

 1. Psicologia cognitiva - Teses. 2. Atenção - Aspectos psicológicos -
 Teses . I. Rossini, Joaquim Carlos. II. Universidade Federal de
 Uberlândia. Programa de Pós-Graduação em Psicologia. III. Título.

CDU: 159.95

Aproximadamente 10 mililitros de tinta preta dando forma a quase 19 mil palavras sobre a celulose processada. Esta é forma física dessa dissertação. Mas o que a construiu é invisível aos olhos. Sem o apoio incondicional dos meus familiares, dos meus amigos e de minha namorada, sem o auxílio de um orientador engajado, ela não passaria de devaneios de uma mente sonhadora. A estes tijolos invisíveis, é que dedico essa dissertação.

Lista de Ilustrações

Figura 1	42
Figura 2	45
Figura 3	47
Figura 4	49
Figura 5	55
Figura 6	57
Figura 7	58
Figura 8	63
Figura 9	65
Figura 10	68
Figura 11	69

Sumário

Lista de Ilustrações	4
Sumário	6
Resumo	8
Abstract	9
1. Introdução	10
1.1 Informação de baixo-nível versus Informação de alto-nível	10
1.2 Estudos Pioneiros	15
1.3 Estudos Posteriores	22
1.4 A informação de baixo-nível como mediadora da informação de alto-nível	34
2. Objetivos	40
3. Experimentos	41
3.1 Experimento 1	41
3.1.1 Participantes	41
3.1.2 Estímulos	41
3.1.3 Procedimentos	44
3.1.4 Resultados e Discussão	45
3.2 Experimento 2	52
3.2.1 Participantes	53
3.2.2 Estímulos	53
3.2.3 Procedimentos	54
3.2.4 Resultados e Discussão	57
3.1 Experimento 3	59

3.3.1 Participantes	61
3.3.2 Estímulos	61
3.3.3 Procedimentos	64
3.3.4 Resultados e Discussão	67
Discussão Geral	70
Bibliografia	79
Aceite do CEP	83

Resumo

De modo geral, as investigações iniciais acerca da atenção seletiva apresentaram uma grande ênfase no estudo do processamento da informação visual de baixo-nível e médio-nível. No entanto, tem crescido o número de estudos sobre a influência da informação de alto-nível na seleção da informação em cenas pseudo-realistas, significativamente mais complexas e contextualizadas. Este estudo foi desenvolvido com o objetivo de estudar a influência da informação semântica na busca visual. Porém, ao contrário de outros no campo, optou-se por utilizar estímulos físicos simples. Nos experimentos, um conjunto de setas paralelamente orientadas e apontadas em uma mesma direção foi usado como elemento semântico, a hipótese era que se a informação das setas guiasse a atenção, a busca seria mais eficiente quando o direcionamento das setas fosse congruente com o hemisfério da localização do alvo. No experimento 1, os resultados demonstraram que quando apresentada em forma de textura com luminância distinta dos elementos relevantes, a informação semântica não influencia a busca. No experimento 2, a congruência da textura com o alvo foi manipulada de forma que em 80% das provas as setas indicavam o hemisfério do alvo. Porém, os resultados novamente mostraram que a informação semântica não influenciou na busca. Por fim, no experimento 3, a ativação da textura foi manipulada através de uma pré-ativação gerada por uma busca anterior. Os resultados sugerem que a pré-ativação das características da textura retardou a busca, porém efeito da informação semântica. Conclui-se, assim, que a informação semântica não influencia na seleção quando usado estímulos físicos simples, sendo descartada quando apresenta baixa similaridade com os estímulos relevantes.

Abstract

Overall, the initial researches about selective attention showed a great focus on the study of low-level and mid-level of information process. Though, there are an increasing number of studies about of high-level information influence on the selection using realistic and complex scenes. This work aims to study the semantic information in visual search. However, unlike others in the field, the method use simple physical stimulus. In the experiments, a set of arrows parallel oriented and pointed to the same direction was used as a semantic element, the hypothesis was that if the semantic information guide the attention, the search would be more efficient when the direction of the arrows were congruent with the hemifield of the target location. In experiment 1, the results showed that when the texture luminance is distinct of relevant elements, semantic information does not influence the search. In experiment 2, the congruency of texture with the target was manipulated so that in 80% of the proofs the arrows indicate the hemifield of the target location. However, the results again showed that semantic information does not influence the search. Finally, in experiment 3, activation of the texture was manipulated by a priming generated by a previous search. The results suggest that ppriming of the textures feautres slowed the search, but there was no significant effect of the semantic information. Thus, was concluded that the semantic information does not influence the selection when used simple physical stimuli, being discarded when it has low similarity with the relevant stimuli.

1.Introdução

1.1 Informação de baixo-nível versus Informação de alto-nível

Enquanto dirigimos nosso carro pelo centro de uma grande cidade somos distraídos por uma grande quantidade de estímulos, como, por exemplo, outros carros, outdoors ou pedestres na rua. Assim, corriqueiramente selecionamos e descartamos estímulos relevantes dentre estímulos irrelevantes contidos em um ambiente saturado de informação. O mesmo acontece com um médico quando procura por um padrão específico em uma imagem radiográfica ou com um controlador de tráfego aéreo que procura por aeronaves em uma determinada rota. Esse processo de seleção da informação é o que caracteriza a **atenção seletiva**.

O estudo da atenção seletiva é comumente realizado através do paradigma da busca visual em que os participantes buscam por um determinado estímulo alvo apresentado em meio a estímulos distratores. O sucesso em localizar o alvo é decorrente da seleção eficaz da informação do ambiente baseada em três principais fontes: a **informação de baixo-nível**, a **informação de médio-nível** e a **informação de alto-nível**. A informação de baixo-nível refere-se às características físicas simples como cor, forma, luminância e movimento. A informação de médio-nível se caracteriza pelo resultado da integração das características físicas simples, envolvendo então a percepção de objetos e a determinação de sua localização espacial. Por último, a informação de alto-nível refere-se à representação visual semântica, dada pela integração dos estímulos à memória de longo prazo, caracterizado pelo reconhecimento da cena, contexto e categorias. A informação de alto-nível é também referida como **informação semântica** (Henderson & Hollingworth, 1999).

Os estudos iniciais sobre a atenção seletiva através da busca visual apresentam uma grande ênfase no processamento da informação de baixo-nível e médio nível,

geralmente utilizando estímulos bastante simples. Esses estudos apontam que quando o alvo tem alguma característica distinta dos distratores, a busca costuma ser rápida e sem esforço, por exemplo, um quadrado azul é rapidamente localizado no meio de quadrados vermelhos. Porém, quando o alvo compartilha características com os distratores a busca costuma demandar mais tempo, por exemplo, na busca de um “T” em meio a “L”s, os dois possuem um seguimento vertical e outro horizontal em comum (Duncan & Humphrey, 1989; Treisman & Gelade, 1980; Wolfe, Cave, & Franzel, 1989).

Uma das teorias usadas na explicação destes fenômenos de busca visual é a **Teoria da Integração das Características (TIC)** (Treisman & Gelade, 1980). Segundo esta, a busca visual seria composta de dois estágios sequenciais e distintos: os processos pré-atentivos e atentivos.

No primeiro estágio, características básicas como cor, forma, orientação seriam processadas e codificadas em lócus separados, os chamados mapas mentais. Cada mapa seria um sistema independente responsável pela codificação de uma característica física específica. Inicialmente não haveria nenhuma comunicação entre os mapas, somente mais tarde com o engajamento da atenção é que essa comunicação seria possível. Todos os estímulos por todo campo visual seriam imediatamente codificados pelo seu mapa correspondente, isto significa um processamento paralelo (vários itens de uma vez), automático e com capacidade ilimitada. Não existiria neste primeiro estágio a codificação da localização das características. Desta forma, o processamento paralelo só pode identificar a presença de uma característica, mas não localizá-la (Treisman & Gelade, 1980).

O segundo estágio inicia-se com o engajamento da atenção. Aqui a atenção atuaria integrando a informação codificada em mapas separados. Assim, por exemplo, uma característica como a cor seria integrada a uma característica como a forma para

dar origem a percepção de um objeto. Através dessa integração das características surgiria o mapa das localizações, sendo então possível determinar a localização de um item. Porém, a atenção só seria capaz de integrar um item por vez, ou seja, trata-se de um processamento serial (um item por vez) de capacidade limitada. Não existiria prioridade no processamento da informação, assim, a seleção dos itens pela atenção seria aleatória, todos os itens do campo visual teriam a mesma probabilidade de serem os próximos a serem processados (Treisman & Gelade, 1980).

Através deste modelo, podemos explicar porque a busca por um quadrado azul em meio a quadrados vermelhos é bem mais eficiente do que a busca de um “T” em meio a “L”s.

O primeiro caso é conhecido como busca por característica. Nesta, o alvo possui uma característica distinta, a cor azul. Assim, basta que o mapa de uma característica relevante seja monitorado para averiguar a presença de alguma ativação. Não é necessário, aqui, a integração da informação. Dessa forma, a busca é mais rápida por envolver somente processamento paralelo.

Uma decorrência disto é que o tempo da busca não aumenta com a adição de mais itens ao tamanho da apresentação. Com o processo paralelo, todos os elementos do campo visual são processados ao mesmo tempo, não importa quantos itens existem, a característica relevante será selecionada imediatamente. Assim, o tempo da busca acaba sendo independente do tamanho da apresentação.

O segundo caso, busca de “T” entre “L”s, é conhecido como busca por conjunção. Nesse o alvo e distratores são formados pelas mesmas características, neste caso, um seguimento vertical unido a um horizontal. Dessa forma, os dois estímulos irão ativar os mesmos mapas no processamento paralelo, não sendo possível determinar

a presença do alvo. Para discriminar “T” do “L” é necessário o engajamento da atenção. Como a atenção só é capaz de integrar um item por vez e o aumento de itens significa um aumento de candidatos a alvo, o tempo da busca cresce com o tamanho da apresentação.

Um dado muito importante é que nos experimentos realizados por Treisman e Gelade (1980), o tempo gasto em busca por conjunções quando o alvo estava ausente é em média o dobro do que quando o alvo está presente. Esta relação 2:1 aponta para um padrão de busca aleatória. Porque para concluir que o alvo está ausente seria necessário verificar todos os itens, mas quando o alvo está presente sua localização é aleatória, podendo tanto ser o primeiro como o último a ser selecionado. Portanto, se o alvo pode ser tanto o primeiro quanto o último item na verificação, ele em média será o item do meio na ordem da verificação. Origina-se desta maneira a relação 2:1 entre provas com alvos ausentes e provas com alvos presentes. Deriva-se disto uma das principais premissas da TIC: as busca por conjunções tem um padrão aleatório de seleção de itens.

Outras teorias, além da TIC, foram propostas para o estudo do processamento da informação na busca visual, tais como: **Teoria da Similaridade** (Duncan & Humphrey, 1989), e a **Teoria da Busca Guiada** (Wolfe, Cave, & Franzel, 1989). Apesar das contribuições importantes de cada uma dessas teorias, todas apresentam uma limitação, abordam quase exclusivamente o processamento da informação de baixo-nível e médio-nível, e não prevêm com eficiência como os estímulos reais e a informação de alto-nível podem influenciar na busca visual.

As primeiras pesquisas sobre a influência da informação visual de alto-nível nos processos cognitivos foram no campo da percepção (Biederman, 1972). Uma preocupação dos pesquisadores da percepção era que, ao contrário dos estímulos

normalmente utilizados em laboratório, as cenas no mundo cotidiano contêm uma quantidade de informação muito maior e complexa. Os modelos clássicos de busca visual prevêem uma busca lenta e serial nestas condições, pois há pouca informação pré-atentiva capaz de guiar a busca, que necessita integrar uma grande quantidade de informação para localizar o alvo. Todavia, apesar da alta quantidade de informação a ser processada, operamos com grande eficácia no ambiente natural. A razão para isto seria que, ao contrário do que ocorre com os estímulos artificiais, a informação visual de uma cena real apresenta uma grande coerência contextual. Em outras palavras, as cenas reais são estruturalmente redundantes de tal forma que dada à visão de determinado local é possível prever quais objetos ali estarão, sua dispersão no espaço e o tamanho relativo entre eles. Por exemplo, ao examinar visualmente uma cozinha podemos prever que haverá uma pia, uma geladeira e um fogão, todos alinhados a parede. Isto acontece porque a cognição opera através de processos altamente adaptativos e auto-organizadores que registram e antecipam as regularidades do ambiente.

Os objetos e ambientes seriam codificados através de representações conhecidas como **esquemas** (Rumelhart & Ortony, 1977). As teorias dos esquemas propõem que o conhecimento é codificado através de representações amplas e genéricas. Essas representações são construídas através da aprendizagem de padrões subjacentes que se repetem na experiência individual. A representação desses padrões gerais permite a antecipação de elementos em um caso particular. Assim, os esquemas agem como sistema de detecção de padrões, que uma vez evocados, guiam a cognição no reconhecimento e compreensão de novos casos. Desta maneira, a visualização de um ambiente ativaria o seu respectivo esquema, no qual estariam codificadas informações como: os objetos mais comuns, sua geral dispersão no espaço e o tamanho relativo entre os objetos desse ambiente. A representação dos objetos e sua dispersão facilitariam o

processamento da informação reduzindo a quantidade de detalhes necessária para identificação de objetos antecipados pelo esquema. Por sua vez, a identificação de objetos que não pertencem ao esquema e são inesperados naquele ambiente necessitaria seria mais onerosa, pois demandaria uma maior análise. Uma vez que os modelos clássicos da atenção prevêm que a informação de alto-nível é acessada tardiamente, a questão que surge é: como o acesso a informação dessas relações contextuais de alto-nível é rápido o suficiente para influenciar a percepção de um objeto em cenas naturais?

1.2 Estudos Pioneiros

Para testar a hipótese da facilitação da percepção pelo contexto, Biederman (1972) delineou um experimento em que os sujeitos observavam fotos de cenas reais em preto e branco em projeções de slides. Cada cena projetada possuía duas versões: uma normal e outra em que a cena era dividida em seis partes e então reorganizada formando uma nova imagem. Na condição de cena reorganizada, as partes da cena ocupavam posições distintas, mas nunca eram rotacionadas. A imagem de setas foi usada para indicar uma área alvo. Em metade das provas, as setas eram apresentadas 300ms antes da cena e na outra metade eram apresentadas 300ms depois da cena. No final era apresentada uma imagem com quatro objetos e o sujeito deveria determinar qual desses estava na área apontado pela seta. Uma segunda versão deste experimento foi realizada em que os objetos eram apresentados antes da foto, devendo o participante responder, ao final da exposição, se um dos objetos estava na área apontada pela seta. Em todos os tratamentos experimentais a versão de cena reorganizada produziu uma taxa maior de respostas incorretas. Este resultado sugere que a percepção do contexto da cena influencia na percepção de objetos singulares. Esta conclusão é possível porque tanto na condição de cena normal quanto na condição de cena reorganizada a estrutura dos

objetos é mantida intacta, o único elemento alterado é sua organização no espaço. Se a percepção do objeto fosse independente da cena, não haveria diferença entre as condições. Porém, na cena reorganizada, o seu o arranjo natural é corrompido, o que dificulta a identificação da cena. Esta alteração impede a ativação do esquema geral, que atuaria normalmente facilitando a percepção dos objetos. Desta forma, há uma tendência maior de erros nas cenas reorganizadas.

Pensando que os objetos não interagem entre si de uma só forma, Biederman, Mezzanotte, & Rabinowitz (1982) descreveram cinco formas em que a informação da relação entre objetos pode ser codificada no esquema geral:

1. **Suporte:** a maioria dos objetos repousa sobre superfícies, somente poucos flutuam no espaço;
2. **Interposição:** os objetos ocupam espaços próprios, não podendo ocupar concomitantemente a mesma posição;
3. **Probabilidade:** certos objetos são mais comumente encontrados em um ambiente do que em outro, por exemplo, é mais fácil encontrar uma panela na cozinha do que um hidrante;
4. **Posição:** os objetos em um ambiente são alocados preferencialmente em certos espaços, por exemplo, é mais comum ver um liquidificador sobre a bancada da cozinha do que no chão;
5. **Tamanho:** os objetos têm tamanhos relativos entre si, por exemplo, é pouco provável que exista no escritório um cortador de papel maior que uma cadeira.

Uma hipótese inicial era que algumas das relações contextuais seriam mais prontamente processadas, o que influenciaria precocemente a percepção. As relações de

Suporte e Interposição seriam as primeiras a serem processadas. Isso porque para perceber que um objeto está flutuando no ar não é preciso identificar qual objeto se trata, basta perceber que o mesmo se encontra sem suporte. Em uma situação hipotética, poderíamos perceber que um objeto esférico flutua no ar sem suporte antes de determinarmos com precisão sua identidade, por exemplo, se este objeto é uma laranja ou uma maçã. O mesmo é válido quando existe uma interposição entre objetos. Nesta situação, não é preciso identificar quais são os objetos para saber que ocupam o mesmo lugar no espaço. Após o processamento das relações de Suporte e Interposição haveria o processamento das relações de Probabilidade e depois Posição. O processamento da Probabilidade é prioritário em relação à Posição, pois, apesar de podermos prever quais objetos uma determinada cena pode conter, a sua localização na cena pode variar. Por exemplo, é possível prever a presença de um fogão em uma cozinha, porém sua posição espacial pode variar significativamente. Finalmente, após esta sequência de processamento, a relação de Tamanho entre os objetos seria avaliada. Esta avaliação ocorreria por último, pois não seria possível perceber alterações no tamanho de um objeto sem ter outros objetos como referência. Um exemplo são os tabuleiros de xadrez gigantes, para perceber seu aspecto incomum é necessário primeiro identificar outros objetos na cena para poder compará-los, e só então, perceber o seu tamanho alterado.

Para testar essa hipótese, Biederman, Mezzanotte e Rabinowitz (1982) delinearam um experimento em que o nome de um estímulo alvo era apresentado visualmente ao participante e permanecia disponível até que o sujeito indicasse que havia memorizado a palavra e estava pronto para iniciar a prova experimental. A prova experimental era iniciada pela apresentação de um sinal de fixação por 500ms, seguido pela apresentação de uma imagem por 150ms. Após a apresentação da imagem, uma seta era apresentada como dica espacial por 500ms. A tarefa do participante era

pressionar uma determinada tecla do computador caso a indicação da seta fosse coerente com a posição do alvo, e outra tecla, caso o alvo não estivesse presente. O alvo era apresentado em duas condições: consistente com a cena ou inconsistente com a cena. Na condição de inconsistência o alvo apresentava de uma à três violações nas relações inter objetos (Suporte, Interposição, Probabilidade, Posição e Tamanho), enquanto na condição de consistência o alvo não apresentava nenhuma violação nas relações inter objetos. A hipótese era que como as relações inter objetos teriam uma seqüência de processamento específica, sendo certas relações processadas antes de outras, a taxa de erros seria diferente dependendo da relação violada.

Os resultados mostraram que os sujeitos cometeram mais erros na condição de alvo inconsistente. Porém, contrariando as hipóteses iniciais, não houve diferença nas violações inter objetos, apresentando taxas semelhantes de erros. Outro ponto, é que as violações apresentaram um efeito aditivo, ou seja, três violações tendiam a produzir mais erros que duas, que por sua vez tendiam a produzir mais erros que uma só. Isso também se refletiu no tempo de reação. Os sujeitos respondiam mais lentamente quando havia uma violação, independente de qual seja, tendo elas também um efeito aditivo. Esses resultados sugerem que a informação semântica da cena incluindo relações entre a cena e objetos podem ser rapidamente acessados, além disso, essa informação é utilizada de forma a facilitar a percepção dos objetos consistentes. Porém, não existe prioridade no processamento das formas de codificação entre as relações entre objetos e cena, já que não houve nenhuma diferença no efeito da resposta dentre as violações.

No entanto, Hollingworth e Henderson (1998) apontaram algumas limitações no delineamento proposto por Biederman *et al.* (1982). Uma limitação importante apontada por eles é que o experimento original não era adequado para controlar a tendência do participante em responder “sim” nas provas controle (em que o alvo estava ausente)

quando a palavra alvo exibida inicialmente era consistente com a cena. Isto pode ter gerado mais alarmes falsos quando o alvo era consistente com a cena e menos na situação de inconsistência. Por exemplo, se o participante procurava por um “cavalo” em uma cena de fazenda, ele tenderia mais a responder “sim” do que quando o procurava em uma cena de uma sala. Esta consistência palavra-alvo ocorria 70% das apresentações em provas de controle, o que pode ter causado um viés significativo. Outra limitação neste experimento era que o participante não tentava detectar o mesmo objeto nas provas em que o alvo estava presente e nas provas controle. Porém, segundo a teoria de detecção do sinal, para calcular eficazmente a sensibilidade de detecção de um sinal é necessário calcular a detecção de um sinal quando ele está presente, e a falsa detecção do mesmo sinal quando ele está ausente (Green & Swets, 1966). Por exemplo, um alvo consistente em uma prova de controle poderia ser “um pote de biscoito” e o objeto indicado pela dica um fogão. A resposta correta refletiria a detecção de um fogão, porém o alarme falso indicaria a falsa detecção de um pote de biscoito. Como o pote de biscoito é menos comum em uma cozinha que um fogão, o alarme falso seria artificialmente baixo, e a estimativa da detecção seria artificialmente alta. Na tentativa de solucionar estas limitações Hollingworth e Henderson (1998) replicaram o experimento, porém, modificando o delineamento para isentá-lo de tais vieses na resposta. Os resultados demonstraram que as respostas dos sujeitos não diferiram entre a condição de objeto consistente com a condição de objeto inconsistente, o que sugere que a percepção do objeto não é facilitada pelo contexto da cena e que os dados encontrados por Biederman *et al.* (1982) foram resultado de um paradigma experimental enviesado (Hollingworth & Henderson, 1998).

Outro estudo seminal foi realizado por Loftus e Mackworth (1978). Estes autores delinearam um experimento em que os participantes viam cenas desenhadas

através de linhas e traços. Nestes desenhos o alvo poderia ser tanto consistente quanto inconsistente com a cena. A informação visual era controlada de tal forma que os alvos e as cenas eram intercambiáveis, por exemplo, um dos alvos poderia ser em um momento um polvo e no outro um trator. Ambos poderiam aparecer tanto em uma cena de fazenda, quanto em uma cena de fundo do mar. Cada cena era apresentada por quatro segundos, sendo os movimentos oculares registrados durante a apresentação. Ao fim da apresentação os participantes eram solicitados a realizar um teste de memória de reconhecimento. Os resultados demonstraram que o número de fixações oculares era significativamente superior sobre os objetos inconsistentes. Estes resultados também mostraram que os objetos inconsistentes tendiam a receber mais primeiras fixações oculares do que os objetos consistentes. O fato de que os objetos inconsistentes tendem a ser fixados nos instantes iniciais da visualização fortalece a hipótese de que uma representação de alto-nível é ativada nos estágios iniciais da visualização influenciando a seleção da informação.

A explicação de que o esquema global seria ativado nos momentos iniciais da visualização é vaga e imprecisa, já que não menciona o meio de apreensão desse esquema. Uma interpretação ainda bastante ampla é que os estímulos naturais teriam alguma propriedade sinérgica especial capaz de ativar rapidamente esquemas gerais em cenas naturais. De Graef, Christiaens e d'Ydewalle (1990) apontaram a fragilidade desta interpretação do modelo de esquema geral e propuseram um modelo mais parcimonioso e integrado aos modelos clássicos da busca visual. Nesta perceptiva, o efeito de facilitação da cena poderia ser mais bem explicado como uma pré-ativação (priming) causada pela relação entre os objetos. Esta pré-ativação seria gerada e sedimentada pela apresentação sucessiva de estímulos semanticamente compatíveis. Nesta situação, o segundo estímulo visualizado seria pré-ativado pelo primeiro. Assim, a percepção visual

de um objeto facilitaria o processamento de outros objetos semanticamente congruentes. A grande diferença desse modelo é que ele não prevê a ativação de um esquema geral da cena nos momentos iniciais da busca, mas uma representação que se consolida a cada momento com o acúmulo de informação, aumentando progressivamente o seu efeito sobre a seleção visual. Assim, não existiria um processo pré-atentivo da informação de alto-nível, apenas um processo contínuo de integração da informação de baixo-nível que resultaria em uma representação de alto-nível. Essa representação aceleraria, por sua vez, o processo de integração da informação de baixo-nível.

Para testar sua hipótese, De Graef *et al.* (1990) delinearam um experimento em que os participantes deveriam buscar por pseudo-objetos (figuras fechadas que não remetem a nenhum objeto existente) em cenas de traços e linhas construídas a partir de fotografias. Durante a busca os movimentos oculares foram registrados. Cada cena possuía cinco versões: uma em que o objeto era consistente com a cena, e quatro em que era inconsistente com algum aspecto da cena (Posição, Probabilidade, Tamanho e Suporte). Os resultados mostraram que durante os primeiros momentos da busca os objetos inconsistentes eram fixados na mesma frequência que os objetos consistentes, ao passo que nos momentos finais da busca, os objetos inconsistentes apresentavam uma maior probabilidade de receberem fixações oculares. Esses dados indicam que a informação de alto-nível não exerce influência nos momentos iniciais da busca, já que não existe diferença entre a probabilidade de fixação entre objetos consistentes e inconsistentes. Porém, em momentos tardios do processamento há uma maior probabilidade de fixação em objetos inconsistentes. Este padrão favorece a hipótese explicativa de De Graef *et al.* (1990) na qual os objetos facilitam a percepção um dos outros através de um processo de pré-ativação (priming). Assim, alocação da atenção se diferencia nos momentos tardios da visualização.

1.3 Estudos Posteriores

Apesar dos trabalhos pioneiros acerca da seleção da informação semântica apresentarem como ponto comum o uso de imagens nas quais um objeto pode ser consistente ou não com a cena, eles apresentam certa distinção quanto a tarefa a ser executada pelos participantes. Enquanto Loftus e Mackworth (1978) utilizam uma tarefa de memória, De Graef *et al.* (1990) utilizavam uma tarefa de busca visual. No primeiro estudo, os dados sugerem que os objetos inconsistentes são fixados primeiramente, o que evidencia, de certa forma, um efeito precoce da informação semântica nos momentos iniciais do processamento da imagem. No segundo estudo, no entanto, os dados sugerem que este efeito semântico ocorre em estágios mais tardios do processamento. Uma questão importante a ser observada é que os resultados observados são provenientes de paradigmas experimentais distintos e, portanto, a discrepância do resultado pode ser atribuída às diferenças nos procedimentos e não ao processamento da informação.

Na tentativa de avançar nesta questão, Henderson, Weeks e Hollingworth (1999) delinearam uma série de experimentos em que foi novamente empregado o uso de imagens nas quais um objeto pode ser consistente ou inconsistente com a cena. As imagens eram construídas de traços e linhas criadas a partir de fotografias reais. Porém em um experimento os sujeitos observavam as imagens com intuito de memorizá-las para um teste futuro de reconhecimento e no outro buscavam por um alvo indicado por uma palavra antes da apresentação da cena. A tarefa dos participantes era pressionar uma determinada tecla caso o alvo estivesse presente e outra caso o alvo estivesse ausente. O objetivo principal deste estudo foi, portanto, investigar em um mesmo procedimento aspectos vinculados tanto a memória quanto ao processo de busca visual.

O resultado do procedimento de memorização de Henderson *et al.* (1999), assim como os resultados obtidos por Loftus e Mackworth (1978), mostraram uma frequência maior de fixação ocular sobre os objetos inconsistentes durante a apresentação. Entretanto, houve uma discrepância em relação aos movimentos oculares. No experimento de Loftus e Mackworth (1978) foi observada uma maior tendência em realizar os primeiros movimento sacádicos em direção ao objeto inconsistente. Porém no experimento de Henderson *et al.* (1999), os primeiros movimento sacádicos tinha a mesma probabilidade de ser direcionado tanto a um objeto consistente, quanto a um objeto inconsistente.

Os resultados do procedimento de busca visual indicaram que os objetos consistentes tendem a ser fixados logo após um movimento sacádico amplo, além de serem localizados mais rapidamente. Já os objetos inconsistentes eram fixados mais tardiamente e necessitavam mais tempo para serem localizados. Este resultado contrasta com os dados obtidos pelo experimento de busca visual de De Graef *et al.* (1990), em que não existia uma maior tendência de fixação ocular sobre os objetos consistentes no início da busca. O fato dos objetos consistentes serem fixados mais rapidamente que os inconsistentes sugerem algum processamento pré-atentivo da informação de alto-nível. Porém, Henderson *et al.* (1990) argumenta que este não é o caso, e o padrão observado, na verdade, é gerado por um maior conspicuidade espacial das cenas. A explicação para isto seria que a informação da identidade da cena somada a análise das características da cena em larga escala, junto com o conhecimento das características do alvo, permite que o alvo seja rapidamente localizado quando consistente.

Este estudo realizado por Henderson *et al.* (1999) levou ao aprimoramento do **Modelo do Mapa de Saliência (Saliency Framework Map)**, este modelo aborda os macro-movimentos oculares durante a visualização de cenas em razão do

processamento visual (Henderson,1992; Morrison, 1984). Segundo este modelo, um mapa de potenciais áreas alvo para movimentos oculares é formado em um estágio precoce do processamento da cena. Nesta situação cada área receberia um grau de ativação diferente formando um mapa de ativações. Assim, a atenção visual seria direcionada para a área de maior ativação no mapa possibilitando a programação do movimento ocular para esta região. Depois de analisada, a ativação da região diminuiria significativamente no mapa. Assim, a atenção seria direcionada para uma nova área de maior ativação. O que determinaria o grau de ativação de uma área seria a informação de baixo-nível, como luminância, contraste, cor, contorno, densidade, e assim por diante. Somente depois das primeiras fixações a informação de alto-nível estaria disponível passando então a contribuir com o aumento da saliência de certas áreas semanticamente relevantes. A partir deste momento, do processamento, as regiões de saliência semântica começam a ter uma ativação maior, o que aumentaria, portanto, a probabilidade de serem fixadas. Assim, a informação de alto-nível não estaria disponível ao processamento pré-atentivo, pois só atuaria no mapa de ativação posteriormente após algumas fixações. É relevante notar como o Modelo do Mapa de Saliência alinha-se com os modelos clássicos da busca visual que propõem um processamento tardio da informação de alto-nível.

Com base neste modelo é possível entender as discrepâncias entre os dados encontrados por Loftus e Mackworth (1978) e Henderson *et al.* (1999). No primeiro, os dados mostraram que os primeiros movimentos oculares eram direcionados ao objeto inconsistente, enquanto no segundo, os primeiros movimentos sacádicos eram aleatórios, não havendo influência da informação semântica. O delineamento adotado por Loftus e Mackworth (1978) foi mais simples e apresentava pouco realismo, os estímulos eram formados por traços mais elementares, além das cenas possuírem menos

objetos. Já os estímulos utilizados por Henderson *et al.* (1999) eram mais complexos, com alto grau de realismo e compostos por mais elementos gráficos. Por esta razão os objetos apresentados por Loftus e Mackworth (1978) eram mais facilmente dissociados do fundo, enquanto que os objetos utilizados por Henderson *et al.* (1999), não. Assim, Henderson *et al.* (1999) argumentaram que os estímulos de Loftus e Mackworth (1978) eram mais salientes quanto as características intrínsecas como luminância, contraste, forma, cor, orientação, gerando áreas de ativação mais proeminentes e mais rapidamente selecionados pela atenção. Por outro lado no experimento de Henderson *et al.* (1999), a baixa saliência dos estímulos não produzia uma ativação robusta dos objetos inconsistente e um não direcionamento da atenção. Neste caso, somente com o acúmulo de informação, os objetos inconsistentes gerariam ativação suficiente para ser selecionada pela atenção.

Rousselet, Fabre-Thorpe e Thorpe (2002) propuseram um experimento capaz de avaliar uma possível capacidade automática para o reconhecimento de objetos familiares. Um aspecto inovador do delineamento proposto por estes autores foi o registro dos potenciais cerebrais relacionados aos eventos, o que permite uma maior precisão no registro temporal do processamento. Isso ocorre uma vez que o registro do tempo de reação motor é uma representação não somente do tempo necessário para o processamento do estímulo, mas também do tempo necessário para a programação e execução da resposta. Neste estudo, a tarefa do participante era pressionar uma determinada tecla quando a imagem apresentada continha um animal. Quando a imagem não continha um animal o participante era instruído a não efetuar nenhuma resposta. As imagens alvos eram de diversos animais em seus ambientes naturais, e as imagens controle eram de florestas, lagos, bosques e montanhas sem a presença de animais. A comparação dos potenciais cerebrais em provas com o alvo e provas controle sugere que

eles diferem sistematicamente em aproximadamente 150ms. No entanto, a resposta dos participantes só ocorre em média com 400ms. Essa diferença nos potenciais cerebrais aos 150ms demonstra que, nessa etapa do processamento, a presença ou ausência do alvo já é identificada. Neste mesmo estudo, foi delineado um segundo experimento em que duas imagens eram apresentadas ao mesmo tempo, assim havia três condições: com ambas as imagens contendo o alvo, com somente uma das imagens contendo o alvo e com nenhuma imagem contendo o alvo. Os resultados das medidas de tempo de reação e potenciais elétricos cerebrais não diferiram significativamente entre as condições de alvo presente. Como o processamento de uma imagem não retarda o processamento da outra, isto sugere um processamento em paralelo da informação de alto-nível, ou seja, um processamento pré-atentivo (Rousselet *et al.*, 2002).

Gordon (2004) também investigou o efeito da informação de alto-nível sobre a alocação da atenção durante os primeiros instantes da visualização. Ele aponta o registro dos movimentos oculares como uma medida enviesada para o estudo da atenção. Isto se deve ao fato do movimento ocular ser uma medida indireta da atenção. Quando usado para entender fenômenos que se prolongam no tempo geralmente é uma medida confiável, porém em fenômenos que ocorrem em um curto espaço de tempo pode existir uma dissociação entre atenção e movimento ocular. Assim, é possível que os estímulos semânticos sejam processados mais prematuramente do que demonstram os movimentos oculares. Neste estudo, como em anteriores, foram usadas imagens de cenas baseados em desenhos com traços e linhas, em metade delas em uma determinada localização estava um objeto consistente com a cena e na outra metade um objeto inconsistente. A diferença dos outros estudos está na tarefa do sujeito. Uma das tarefas era identificar um estímulo alvo (“%” ou “&”) que era apresentado após a imagem da cena. O participante deveria pressionar um determinado botão no mouse de acordo com

o alvo apresentado. No mesmo experimento, uma segunda tarefa concomitante era identificar o nome de um objeto presente na cena em um conjunto de alternativas. O objetivo desse procedimento era investigar se os recursos atentos eram mobilizados pelos estímulos inconsistentes e consistentes em relação à cena. A hipótese básica deste procedimento era que os sujeitos apresentariam tempos de reação menor quando o estímulo alvo era apresentado em uma posição previamente ocupada por objetos que estariam mobilizando recursos atentos.

Os resultados demonstraram que a resposta foi mais rápida e precisa quando o alvo era apresentado depois de um intervalo de 40 à 70ms, em uma posição previamente ocupada por um objeto consistente. Por outro lado, depois de 150ms, a resposta dos participantes era mais rápida e precisa quando o alvo era apresentado em uma posição previamente ocupada por um objeto inconsistente. Desta forma, estes dados sugerem que a informação semântica influencia a atenção, mesmo nos seus momentos iniciais do processamento. Na segunda tarefa concomitante, em que os participantes deveriam identificar o nome de um objeto presente na cena em um conjunto de alternativas apresentadas após que a primeira tarefa fosse concluída, os participantes apresentaram uma alta taxa de erros. Uma possível interpretação desses resultados é que a atenção utiliza dados locais e gerais para determinar o contexto da cena em momentos precoces do processamento. Assim, os recursos atentos são alocados em objetos consistentes para confirmar as hipóteses iniciais acerca de um contexto. Por outro lado, com o acúmulo de informação ao longo do processamento, os objetos inconsistentes geram um conflito com a hipótese inicial de coerência da cena. O fato dos participantes apresentarem pouca precisão no reconhecimento do objeto na tarefa de memória pode indicar que, apesar de haver uma identificação visual primitiva que faz com que a atenção seja alocada inicialmente em objetos consistentes e depois em inconsistentes, não

há informação suficiente para a plena integração do objeto. Uma segunda interpretação seria que a tarefa de identificação do alvo produz uma interferência que dificultaria a retenção do objeto na memória imediata.

Os achados de Gordon (2004) e Rousselet *et al.* (2002) corroboram a **Teoria de Memória Conceitual de Curto-Prazo (MCCP)** (Potter, 1993; 1999). Esta teoria sugere que o processamento seletivo da informação semântica é um processo mais rápido do que normalmente se pressupõem. A hipótese proposta é que os estímulos semânticos são rapidamente identificados e seu significado é ativado e mantido brevemente na memória conceitual de curto-prazo. Esta representação inicial é genérica e imprecisa, sendo utilizada na elaboração de uma estrutura de representação compreensiva capaz de “guiar” o processo de seleção da informação. Porém, quando essa informação não é estruturada ela esvanece rapidamente. Esta forma de memória difere da memória de trabalho e de longo prazo, pois prediz que o processamento da informação semântica ocorre de maneira automática com pouca ou nenhuma consciência.

Interessados na relação da informação semântica com a saliência física dos estímulos, Underwood e Foulsham (2006) propuseram um estudo composto por dois experimentos. Assim como no estudo de Henderson *et al.* (1999), um experimento apresentou uma tarefa de memorização e outro uma tarefa de busca visual. Os estímulos utilizados neste estudo foram fotografias de ambientes internos. Em cada cena foram apresentados dois objetos alvos, um com alta saliência física e outro com baixa. Além disso, cada dupla de objetos era apresentado em duas condições. Em uma das condições um dos objetos era consistente com a cena, enquanto que o outro objeto era inconsistente. Na outra condição a relação dos objetos com a cena se invertia. Os resultados mostraram que durante o teste de memorização os objetos com alta saliência

física tinham maior probabilidade de receberem fixações iniciais. Porém, os objetos inconsistentes tinham a maior duração de fixações independente da sua saliência física. Já na tarefa de busca visual, o tempo de latência era menor na condição de objetos consistentes em relação aos objetos inconsistentes. Os objetos consistentes tinham maior probabilidade de receber fixações iniciais independente da saliência física.

A maior influência da saliência física na tarefa de memória em relação à tarefa de busca visual levou Underwood e Foulsham (2006) a concluir que a atenção se comporta de forma diferenciada em cada tarefa. Esta diferença atensiva, segundo os pesquisadores, é gerada pelo nível de direcionamento cognitivo da tarefa. Na tarefa de memorização, o participante não possuía informação prévia sobre o estímulo apresentado, assim a ativação de áreas no mapa seria influenciada principalmente pelas características físicas. No entanto, na tarefa de busca visual existe o conhecimento de um alvo a ser localizado, esse conhecimento influencia diretamente na ativação das áreas no mapa. Desta maneira, em tarefas que existe pouco direcionamento cognitivo, a saliência física tem mais influência sobre atenção, enquanto que em tarefas em que há mais direcionamento cognitivo ocorre uma maior influência da informação semântica. A idéia é que a atenção teria a informação de baixo-nível como sistema básico para guiar a seleção dos estímulos. Porém, quando elementos de alto-nível são importantes para tarefa, a atenção seria capaz de rapidamente selecionar informação do contexto possibilitando a criação de um mapa de ativação capaz de guiar a busca. A hipótese central, e mais importante desse raciocínio, é que o processamento da informação de alto-nível seria pré-atentivo e geraria uma representação capaz de guiar a busca nos seus estágios iniciais.

Em outro experimento Underwood, Templeman, Lamming, e Foulshan (2008), apresentaram aos participantes duas imagens simultâneas dispostas lado a lado. Cabia

aos participantes determinar, o mais rápido possível, a existência de alguma diferença entre as imagens. As imagens eram fotografias de ambientes internos como escritório, sala, quarto, cozinha ou lavanderia. Metade dos pares de fotografias era diferente entre si, enquanto que a outra metade, idêntica. Nas fotografias discrepantes o objeto diferente foi manipulado em quatro condições: consistência com a cena, inconsistência com a cena, alta saliência física e baixa saliência física. Os resultados desse procedimento mostraram que os objetos inconsistentes com a cena eram fixados anteriormente aos objetos consistentes, mesmo quando comparado a condição em que o objeto inconsistente apresentava baixa saliência e o objeto consistente apresentava alta saliência. Este resultado sugere que o fator saliência desempenha um papel secundário no controle das fixações oculares. Na interpretação dos pesquisadores, os objetos inconsistentes foram reconhecidos através de uma representação parcial que incluía forma, cor, orientação, e outras características físicas que leva o objeto a ser identificado como não pertencente à cena.

Outra forma que o contexto de uma cena pode facilitar a localização do alvo é a chamado **dica contextual** (Chun & Jiang, 1998). A hipótese da dica contextual prediz que o layout geral de dispersão do alvo e dos distratores pode ser aprendido de maneira implícita e facilita a localização do alvo em buscas futuras. Uma vez que os ambientes naturais são estruturalmente redundantes, as relações subjacentes entre os objetos podem ser assimiladas, e, então, usadas para ativar o local de maior probabilidade de localização do alvo. A diferença da dica contextual, é que nesse modelo, a identidade do alvo ou distratores não importa, nem sua natureza semântica, o que é aprendido é a sua dispersão e organização no espaço. Para testar essa hipótese Chun e Jiang (1998) realizaram um experimento onde o alvo era um “T” que podia estar rotacionado para direita ou para esquerda apresentados em meio a letras “L”s, também rotacionadas,

utilizadas como distratores. A tarefa dos participantes era identificar se a rotação em 90° do “T” era para direita ou para esquerda. Foram apresentados 12 tipos de layout que se repetiam durante o experimento, e outros layouts que jamais se repetiam. Os resultados mostraram que os participantes eram mais eficientes em localizar alvos em layouts repetidos do que em novos arranjos. O interessante é que os participantes, quando questionados, afirmaram não perceber que os layouts se repetiam. Isso sugere que não houve uma memorização explícita dos arranjos demonstrando que a estrutura de dispersão geral produz áreas espaciais de maior ativação que são memorizadas de forma implícita.

Uma crítica que pode ser feita ao modelo de Chun e Jiang (1998) é sua falta de realismo e validação ecológica, já que os estímulos usados foram simplesmente letras. Fica, então, a pergunta: o display dos objetos em uma ambiente natural também é aprendido de forma implícita e pode guiar buscas futuras ou os estímulos naturais são demasiadamente complexos e a dica de contexto só age sobre estímulos simples? A redundância estrutural dos ambientes naturais e a relativa estabilidade das relações entre os objetos e o seu contexto permitem supor que a dica de contexto apresente um efeito, não só com estímulos artificiais, mas também em ambientes com estímulos naturais. Investigando essa possibilidade, Brockmole e Henderson (2006) delinearam um experimento em que os participantes buscavam por uma determinada letra em fotografias reais. Foram também usados dois tipos de layouts, novos e repetidos. O tempo de busca foi constante nas buscas em layout novo, porém na busca em layout repetido foi observado uma diminuição temporal para a detecção do alvo ao longo das repetições. A velocidade da busca em layouts repetidos chegou a ser até cinco vezes mais rápidas do que a com layouts novos. Este fato sugere que o efeito da dica contextual também é válido em contextos com estímulos naturais. Em um segundo

experimento foi requisitado aos participantes para discriminarem, em uma dada sequência de cenas, aquelas que eles reconheciam do experimento anterior e quais eram novas. O resultado obtido, ao contrário do esperado com base no estudo de Chin e Jiang (1998), mostrou que os participantes tinham consciência da repetição dos layouts. Uma possível interpretação para esta capacidade de memorização explícita dos layouts com imagens reais pode ser atribuída a uma maior facilidade para a codificação de cenas naturais em comparação à memorização de dispersões aleatórias de letras. Possivelmente, na busca em ambientes naturais, tanto a memória explícita quanto a memória implícita devem agir guiando a busca.

Uma pergunta de grande importância no modelo de dica contextual é saber qual aspecto do layout é crítico para guiar a busca. Em um estudo utilizando estímulos artificiais foi observado que o posicionamento do alvo é associado mais ao layout dos distratores próximos do que ao arranjo global da cena (Jiang & Wagner, 2004). Assim, a associação entre dispersão de distratores e alvo pode ser transferida a um novo contexto, contanto que a configuração local do posicionamento do alvo permanecesse inalterada. Porém quando a configuração global permanece inalterada e a configuração local é modificada a aprendizagem não é transferida, não ocorrendo facilitação na busca. Novamente a questão que emerge deste experimento é se este padrão de associação de configuração global e local também ocorre em cenas naturais. Brockmole, Castelhamo e Henderson (2006) fizeram uso de uma configuração global e local de cenas naturais no aprendizado e transferência do layout em uma tarefa de busca visual. Neste experimento os participantes buscavam por uma letra aleatória inserida em uma cena realista. Os estímulos eram cenas realistas tridimensionais criadas com o uso de um software de design de interiores. Novamente a repetição de cenas foi manipulada, havendo cenas novas e cenas já vistas. A cena nunca era rerepresentada completamente inalterada na

condição de cena repetida, sempre a configuração local (elementos próximos ao alvo) ou a configuração global eram alteradas. Por exemplo, em uma cena em que o alvo estava localizado sobre uma mesa de centro ao lado de objetos decorativos, envolvida por uma decoração de biblioteca com estante e livros, a configuração local se caracteriza pela mesa de centro com seus objetos decorativos e a configuração global a decoração de biblioteca com estantes e livros. Em uma condição de cena repetida em que a configuração global é alterada, os arredores podem ser substituídos por objetos de uma sala de estar enquanto a mesa de centro continua intacta. Da mesma, em uma condição de cena repetida com configuração local alterada a mesa de centro onde o alvo está localizado pode ser substituída por uma escrivaninha de madeira, enquanto a decoração de biblioteca com estantes e livros permanece intacta.

Os resultados mostraram que tanto a configuração local quanto a configuração geral eram aprendidos e transferidos para novos contextos. Porém, nas buscas em que a configuração geral mantinha-se inalterada, o desempenho dos participantes era sensivelmente mais rápido do que o demonstrado em configuração local repetida. Esse resultado contrasta em certos aspectos com os dados obtidos com o uso de estímulos artificiais. Nesta condição, além da configuração global ser aprendida e guiar buscas futuras, o que não acontece com estímulos artificiais, ela tem uma influência muito maior do que a configuração local. A explicação para isso está possivelmente nas diferenças entre as características dos estímulos artificiais e naturais. Primeiramente, os estímulos naturais propiciam uma aprendizagem explícita, possibilitando que a configuração geral seja mais facilmente codificada na memória do que a simples dispersões aleatórias de letras em contextos artificiais. Esta aprendizagem explícita pode gerar um efeito mais robusto top-down na busca em cenas locais já conhecidas. Enquanto que a configuração local pode ser baseada em algum tipo de memória

implícita que gera um padrão de ativação mais tênue. Além disso, os elementos realistas gerais de uma cena possuem restrições espaciais que facilitam, por si só, a busca visual.

1.4 A informação de baixo-nível como mediadora da informação de alto-nível

Alguns dos estudos anteriormente citados (Gordon, 2004; Rousselet, *et al.*, 2002; Underwood e Foulsham, 2006) apresentam resultados que sugerem que as representações de alto-nível de um objeto ou de um conjunto de objetos podem ser processadas rapidamente. Isto a princípio parece contrariar o princípio de que o processamento de alto-nível de um estímulo é efetuado, necessariamente, com o engajamento dos recursos da atenção que proporciona a integração das características dos objetos. A fim de investigar essa questão, Evans e Treisman (2005) propuseram a hipótese que os participantes são capazes de detectar um conjunto de características, usando-as para discriminar entre cenas com a presença ou ausência dos alvos, sem que seja necessária a sua completa identificação. As características não precisam ser simples como cor, forma, luminância, etc, estudos já demonstraram que o sistema cognitivo pode aprender características de complexidade moderada, e utilizá-las em uma busca em paralelo (Enns & Resink, 1991; Ramachandran, 1988; Treisman, 1988). Treisman e Gelade (1980) definiram inicialmente característica como um determinado valor em uma dimensão perceptual, por exemplo, laranja seria uma característica na dimensão cor ou triângulo seria uma característica na dimensão forma. Uma característica de complexidade moderada seria o conjunto de algumas características simples e genéricas que não constituem um objeto per se. Assim, características como a forma de um bico de um pássaro ou suas asas, poderiam ser usadas para detectar o estímulo “ave”, da mesma maneira, um conjunto de rodas e textura metálica pode ser utilizado na detecção do estímulo “carro”, sem que seja necessário identificar qual ave ou qual carro se trata. A idéia central é que a detecção de certas características pode mediar a discriminação de

elementos de alto-nível, sem que seja necessário a identificação completa destes. Na busca pelo alvo, características pertinentes a uma categoria buscada, como animais, seriam registradas, o que ativaria conexões relacionadas em uma rede semântica. Tal ativação seria suficiente para gerar uma resposta de detecção, porém a informação não seria suficiente para a identificação completa do objeto.

Um conjunto de experimentos foi realizado para testar essa hipótese. O paradigma utilizado foi o de apresentação visual rápida e serial (rapid serial visual presentation) em que um conjunto de imagens é apresentado rapidamente (75ms por imagem). Este paradigma de apresentação de estímulos foi usado com o objetivo de evitar a integração das características uma vez que o tempo disponível para esse processamento é reduzido. Os participantes foram requisitados a apertar uma determinada tecla assim que detectassem o alvo, o qual poderia ser um animal ou veículo, sendo que para controle existia um conjunto de imagens com alvo ausente. Uma vez pressionada a tecla, a apresentação era interrompida, e então, era requisitado ao participante digitar qualquer informação que eles lembrassem sobre o alvo (como grupo superordenado, como mamíferos, aves, répteis, anfíbios e peixes, alguma característica, como cauda, bicos, asas, quatro pernas e assim por diante) e depois determinar sua localização (direito, esquerdo ou no centro da imagem). Foram utilizadas duas condições de distratores, uma com o conjunto de imagens contendo humanos e na outra contendo plantas.

Os resultados mostraram que quando os distratores são humanos a detecção de animais fica prejudicada, porém a de veículos não sofre influencia. A razão disso pode ser porque os animais e os humanos compartilham muitas características básicas, como olhos, boca, pernas, etc. Assim, a presença de humanos restringiria o número de características discriminativas. O mesmo não acontece com plantas, apesar de

freqüentemente as imagens compartilhem características de cenas de fundo com animais, como imagens de campos e florestas, as características entre plantas e animais são suficientemente distintas para gerarem uma detecção eficaz. Caso a identificação fosse completa os participantes seriam capazes de detectar animais com a mesma eficiência observada na condição com distratores humanos e com plantas. Portanto, a interferência na detecção de animais por distratores humanos é uma evidência favorável a idéia que as características de complexidade moderada mediam a discriminação entre alvos.

Os resultados do questionamento posterior a prova revelou que os participantes eram eficazes em determinar o grupo, porém retinham informações pouco específicas sobre o alvo. Além disso, a resposta de localização não foi mais precisa do que se fosse gerado pelo acaso. Se houvesse uma identificação completa do alvo, os participantes deveriam ser mais específicos na determinação e localização do alvo. O sucesso dos participantes em determinar o grupo superordenado advém da discriminação de características de complexidade moderada, como “bicos” e “asas” para aves, “pelos” para mamíferos, e “pele lisa” para répteis, entre outras características. No entanto, como não há a integração da informação, não é possível para o participante realizar uma discriminação mais detalhada, ou mesmo determinar a localização do alvo. Isto de certa forma confirma a hipótese inicial que as características físicas modulam a detecção dos alvos, sem a necessidade de integração da informação e a sua identificação explícita.

Investindo na idéia que características físicas podem agir como mediadoras na detecção de alvos pertencentes a uma determinada categoria, Levin, Takare, Miner & Keil (2001) realizaram um estudo investigando características físicas utilizadas na discriminação entre artefatos e animais. Em um primeiro experimento, os participantes eram orientados a buscar por um alvo não especificado, cuja única informação

disponível ao participante era a categoria a que pertencia (a imagem de um artefato ou de um animal). O alvo era apresentado entre distratores de categoria oposta (artefatos x animais). Como controle, em metade das provas, o alvo estava presente e na outra ausente. Os resultados mostraram que, para as condições de alvo presente, o custo temporal por item foi de 5.5 ms/item na busca de artefatos entre animais e 16 ms/item na busca de animais entre artefatos. Os resultados demonstram uma grande eficiência na localização do alvo, por exemplo, em uma busca onde não existe nenhuma característica discriminadora com estímulos artificiais em que o alvo é “2” e os distratores são “5” o custo por item é 40 ms/item. Isto sugere que alguma característica comum as categorias está sendo processada de forma a discriminar entre as categorias. Os artefatos possuem traços mais retilíneos quando comparados aos animais, uma hipótese é que esta seja uma das principais características que guia a busca e discriminação entre artefatos e animais. Para testá-la foi feito a análise dos elementos retilíneos dos itens, através de um software específico, gerando, assim, um escore relacionado à quantidade de elementos retilíneos. Em seguida, tal valor foi comparado com a eficiência da localização do alvo. O resultado demonstrou que quanto maior o escore de elementos retilíneos do artefato, mais facilmente este era localizado, enquanto menor o escore do animal, mais dificilmente este era localizado.

Outro estudo da influência de informação de baixo-nível na detecção de elementos de alto nível é o modelo computacional de reconhecimento de cenas reais de Oliva e Torralba (2001). Os autores propuseram que a cena pode ser detectada de maneira holística, sem a necessidade de reconhecimento dos objetos. O reconhecimento da cena aconteceria através da codificação de elementos da configuração espacial global, sem necessidade do processamento de detalhes. A idéia central desta proposta é que ao invés de ver a cena como um conjunto de objetos, a cena per se, é um objeto com

características físicas particulares. Assim, como animais e artefatos possuem características peculiares, cada categoria da cena possui características próprias e similares entre si. Oliva e Torralba (2001) sustentam que estas características, denominadas como “envelope espacial”, são detectadas de maneira pré-atentiva, sem necessidade de segmentação da imagem. O envelope espacial é caracterizado pelas fronteiras e limites que constituem a imagem, como o chão, as paredes, as seções e as elevações do cenário. Por exemplo, auto-estradas geralmente são formadas por um ambiente aberto com longas superfícies horizontais preenchidas com elementos convexos (veículos), enquanto florestas apresentam ambiente fechado com um fundo horizontal bem estruturado (árvores) conectado por uma textura horizontal (vegetação rasteira). Assim, o envelope espacial é representado pelas relações entre linhas externas superficiais e suas propriedades internas de textura.

Para determinar quais seriam os elementos constitutivos e próprios que formaria o envelope espacial de cada categoria da cena, foi solicitado a 17 observadores que dividissem 81 imagens em dois grupos, depois cada grupo em dois novos grupos, até que formasse oito grupos distintos. Os participantes foram orientados a não classificar os objetos pelo seu significado semântico, mas sim com base em elementos físicos estruturais gerais. Depois de realizado a divisão, foi requisitada ao participante que relatasse qual tinha sido o critério utilizado para as divisões. A partir dessas respostas foi gerada uma taxonomia com oito categorias, que por sua vez serviu de base para a criação de cinco propriedades categóricas inerentes ao envelope espacial: Grau de Naturalidade; Grau de Abertura; Grau de Expansão, Grau de Aspereza e Grau de Irregularidade. Em seguida foi construído um modelo computacional capaz de identificar essas categorias. A capacidade de categorização das cenas pelo modelo computacional foi testada comparando o resultado de categorizações com

categorizações humanas. Os resultados demonstraram um alto grau de semelhança entre as respostas, o que sugere que a discriminação de cenas pode ser realizada sem a necessidade do reconhecimento detalhado dos objetos nela contidos.

Apesar de existirem dados favoráveis a hipótese da rápida ativação do esquema geral da cena nos momentos iniciais da visualização, o mecanismo pelo qual isso é possível ainda é uma questão controversa. É paradoxal como a representação de uma cena possa ser formada antes que os elementos que a constituem sejam identificados. Se a cena se caracteriza pelos elementos que a forma, como ela pode ser identificada sem que seus elementos sejam reconhecidos antes, é como se fosse possível comer um bolo antes que seus ingredientes fossem misturados e cozidos. No entanto, o estudo conduzido por Oliva e Torralba (2001) sugere uma solução eficaz para o problema. Por meio do envelope espacial os aspectos físicos holísticos da cena podem ser processados pré-atentivamente, originando uma representação da cena capaz de guiar a atenção. O seu modelo computacional fornece dados robustos de que tal processamento é possível e esclarece o seu mecanismo de funcionamento.

2. Objetivos

Este estudo foi desenvolvido com o objetivo de estudar a influência da informação semântica na busca visual. Os estudos que investigam a informação semântica costumam utilizar estímulos pseudo-realistas com características físicas complexas. Porém, no presente estudo optou-se por utilizar estímulos físicos simples. A razão para isso é investigar se a influência da informação semântica, presente em estímulos de cenas naturais e complexas, pode ser replicada por estímulos físicos simples.

O trabalho tem um caráter exploratório, objetivou-se avaliar se o efeito da informação semântica não é uma propriedade exclusiva de estímulos com características realistas derivada do esquema geral da cena, e sim, uma propriedade do processamento geral da informação de alto-nível presente também em estímulos mais simples. Além disso, com o uso de estímulos simples é possível reduzir e controlar melhor as variáveis, podendo, assim, estudar mais cuidadosamente o papel da informação semântica na busca visual. Por exemplo, uma das grandes dificuldades com estímulos pseudo-realistas é definir o tamanho da apresentação, estimar quais itens são selecionados para uma busca atensiva e quais são descartados como fundo ou textura, enquanto com estímulos simples já é possível definir com precisão o tamanho da apresentação e elementos de fundo.

3. Experimentos

3.1 Experimento 1

Foi usada uma textura composta por imagens de setas paralelas apontando em uma mesma direção como estímulo semântico. A textura era sempre apresentada com um padrão de luminância significativamente diferente dos estímulos relevantes. A hipótese é que se a informação semântica relacionada à textura a guiar a atenção, os recursos atencionais irão se alocar no lado indicado pelas setas. Assim, quando o direcionamento da textura e a posição do alvo forem congruentes (por exemplo, setas apontam para direita e alvo está no hemisfério direito) o alvo será detectado mais facilmente, gerando um menor tempo de reação e menor taxa de erros. Já, quando o direcionamento da textura e a posição do alvo forem incongruentes, o alvo será detectado mais dificilmente, gerando um maior tempo de reação e maior taxa de erros

3.1.1. Participantes

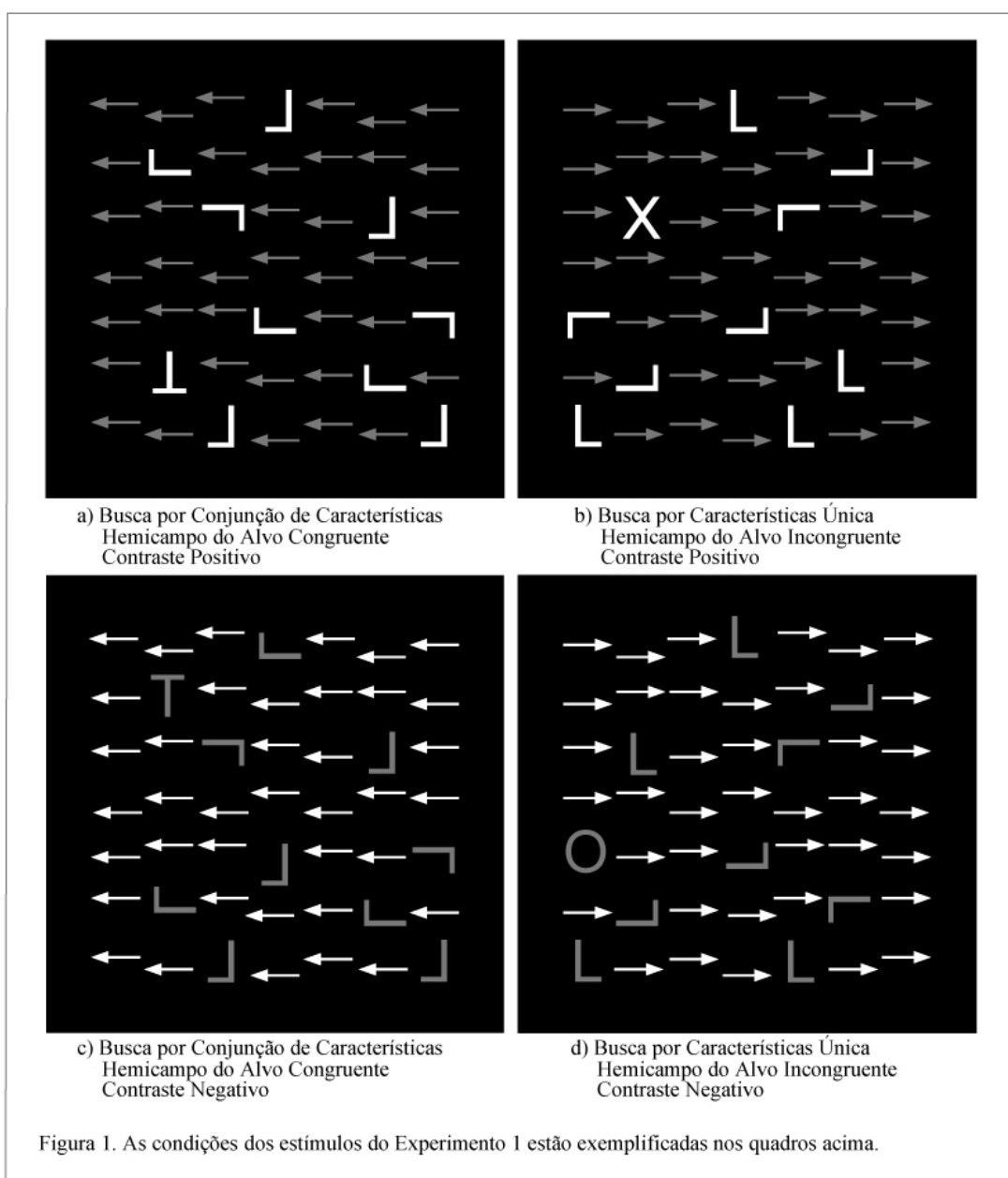
O Experimento 1 foi realizado por 8 voluntários, sendo 2 homens e 6 mulheres, todos destros e apresentando visão normal ou corrigida. A idade média do grupo foi igual à 24 anos ($DP = 1,5$). Os participantes foram esclarecidos acerca das características do experimento, bem como sobre a preservação do sigilo da sua identidade. Após todos os esclarecimentos necessários, e imediatamente antes do início da tarefa experimental, todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido.

3.1.2. Estímulos

Os estímulos visuais foram apresentados em um monitor de 17 polegadas de tubo catódico, configurado em uma resolução de 800x600 pixels com 24 bit de profundidade de cor. A apresentação dos estímulos foi efetuada através do aplicativo E-prime 2.0 (Psychology Software Tools, Inc) instalado em um microcomputador. Para

obtenção dos valores da luminância dos estímulos foi usado um fotômetro GigaHertz HCT- 99D ajustado para o registro em cd/m^2 .

Havia duas condições de busca, uma por **característica única** e a outra por **conjunção de características**. O estímulo alvo na tarefa de busca por característica única consistia em um “X” (1,7° de largura e 2° de altura) em metade das provas e um “O” (1,7° de largura e 2° de altura) em outra metade. Os estímulos distratores nesta condição eram “L”s (1,5° de largura e 2° de altura) rotacionados em 0°, 90°, 180° e



270° e espelhados horizontal e verticalmente. Na tarefa de busca por conjunção de características, o estímulo alvo consistia em um “T” (1,5° de largura e 2° de altura) normalmente orientado em metade das provas e um “T” (1,5° de largura e 2° de altura) rotacionado em 180° na outra metade. Os estímulos distratores consistiam em “L”s (1,5° de largura e 2° de altura) rotacionados em 0°, 90°, 180° e 270° e espelhados horizontal e verticalmente. As texturas usadas consistiam em imagens de setas (com haste com 2,3° de largura e ponta com 0,8° de altura) paralelamente orientadas e apontando para um mesmo lado. Os estímulos da busca eram sempre apresentados no centro da tela em uma matriz com 7 linhas e 7 colunas (com lados medindo 16°), compondo 49 áreas, que quando não estavam ocupadas por um estímulo relevante, eram preenchidas pela textura.

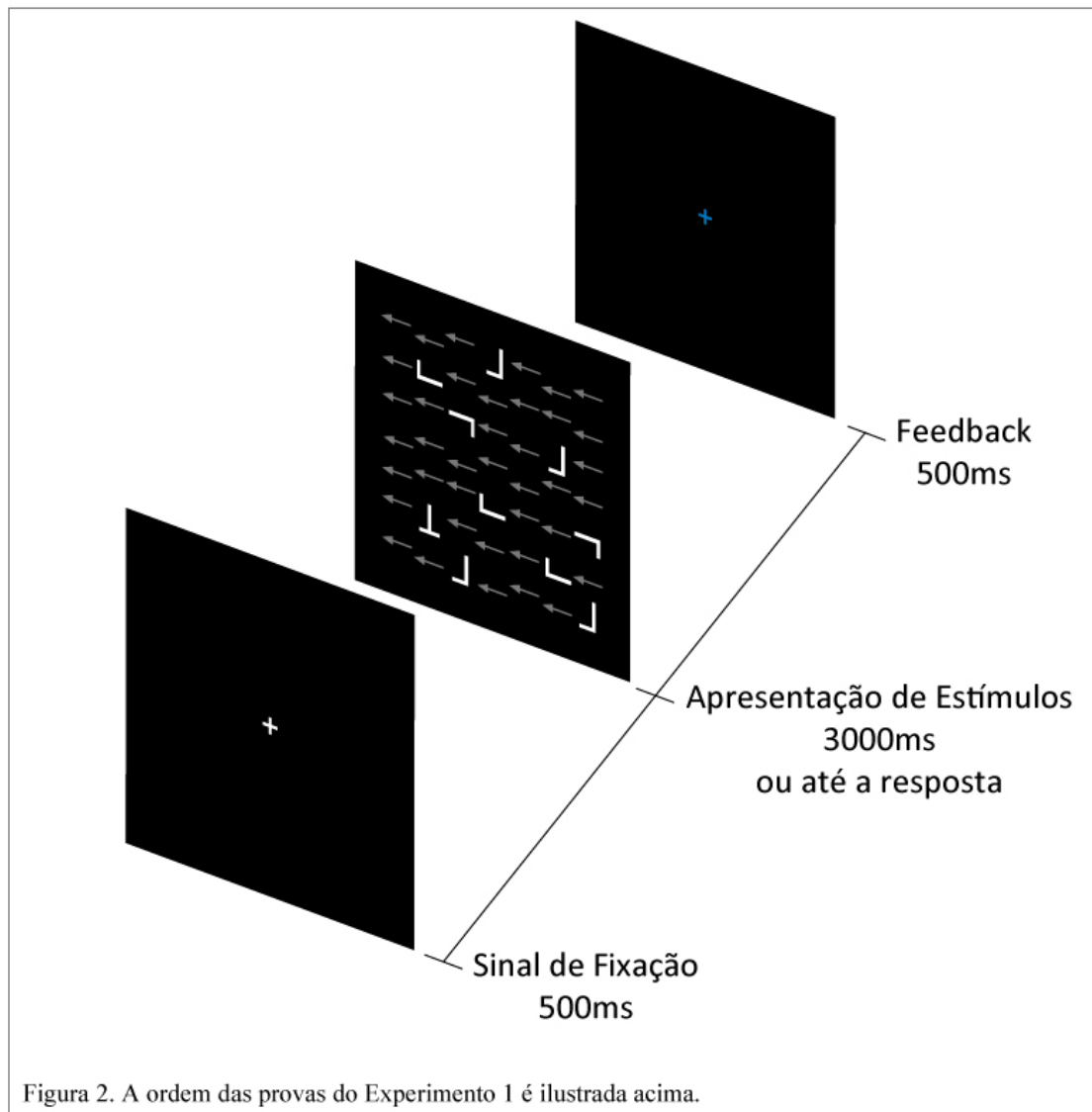
Os estímulos apresentavam duas condições de cor e luminância, uma nomeada de **contraste positivo** e a outra de **contrate negativo**. Na condição de contraste positivo, os estímulos relevantes (alvo e distratores) eram apresentados na cor branca, enquanto a textura era apresentada na cor cinza . Na condição de contraste negativo, os estímulos relevantes (alvo e distratores) eram apresentados na cor cinza, enquanto a textura era apresentada em cor branca. Em todas as condições o fundo da cena era apresentado em negro. Os estímulos apresentados na cor branca possuíam 6,9 cd/m² de luminância, os apresentados na cor cinza possuíam 4,1 cd/m² de luminância, enquanto o fundo negro possuía 1,85 cd/m² luminância. O contraste entre os estímulos relevantes e a textura foi de aproximadamente 25%, calculado através da fórmula de Michelson ($C = \frac{L_r - L_t}{L_r + L_t}$, onde: C =contraste da luminância entre os estímulos, L_r = luminância dos estímulos relevantes e L_t =luminância dos elementos da textura.) A Figura 1 apresenta um exemplo esquemático dos estímulos utilizados.

3.1.3. Procedimentos

Os participantes foram acomodados em uma cadeira confortável a uma distância de 57 cm do monitor, em uma sala com luminosidade controlada. A tarefa dos participantes era procurar por um determinado alvo e responder, o mais preciso e rápido possível, pressionando a tecla correspondente em um teclado numérico. As teclas a serem pressionadas variavam de acordo com a condição dos estímulos relevantes, sendo na condição de busca por características “1” para “X” e “2” para “O”, e na condição de busca por conjunção “1” para “T” em orientação normal e “2” para “T” rotacionado em 180°.

Cada participante realizou duas sessões experimentais em dias distintos. Cada sessão experimental era composta por dois blocos de 264 provas apresentados com um intervalo de dez minutos. Na primeira sessão, foram apresentados os estímulos na condição de contraste negativo, enquanto na segunda sessão, os estímulos foram apresentados na condição de contraste positivo. Em ambas as sessões, a condição de estímulos relevantes era de busca por característica no primeiro bloco e de busca por conjunção no segundo bloco.

No início da sessão eram apresentadas instruções na tela, orientando o participante acerca da natureza da tarefa. Imediatamente após a leitura das instruções e possíveis esclarecimentos, os participantes realizaram um bloco de treino de vinte provas, nas quais as respostas foram descartadas. Cada prova era iniciada com a apresentação de um sinal de fixação ocular “+” (0,8° de largura e comprimento) no centro da tela por 500ms. Após esse intervalo, o sinal de fixação era substituído pela apresentação dos estímulos de busca que permanecia na tela por 3s ou até que o participante respondia. Caso a resposta fosse correta, era apresentado como feedback o sinal de fixação “+” no centro da tela na cor azul, caso fosse incorreta era apresentado



na cor era vermelha. Quando a resposta tardava por mais que 3s, um sinal de interrogação na cor verde (“?” com 0,8° de largura e comprimento) era apresentado no centro da tela. Esses sinais de feedback perduram por 500ms e a prova era encerrada. A ordem da prova é ilustrada pela Figura 2.

Nos blocos experimentais, 3 fatores foram manipulados e igualmente distribuídos entre as provas:

1. **Tipo de alvo:** Os alvos se diferenciarão de acordo com a condição de estímulos relevantes, porém sempre existiam dois tipos (“X” e “O”, para

condição de busca por características, e “T” e “T” rotacionado 180°, para busca por conjunção);

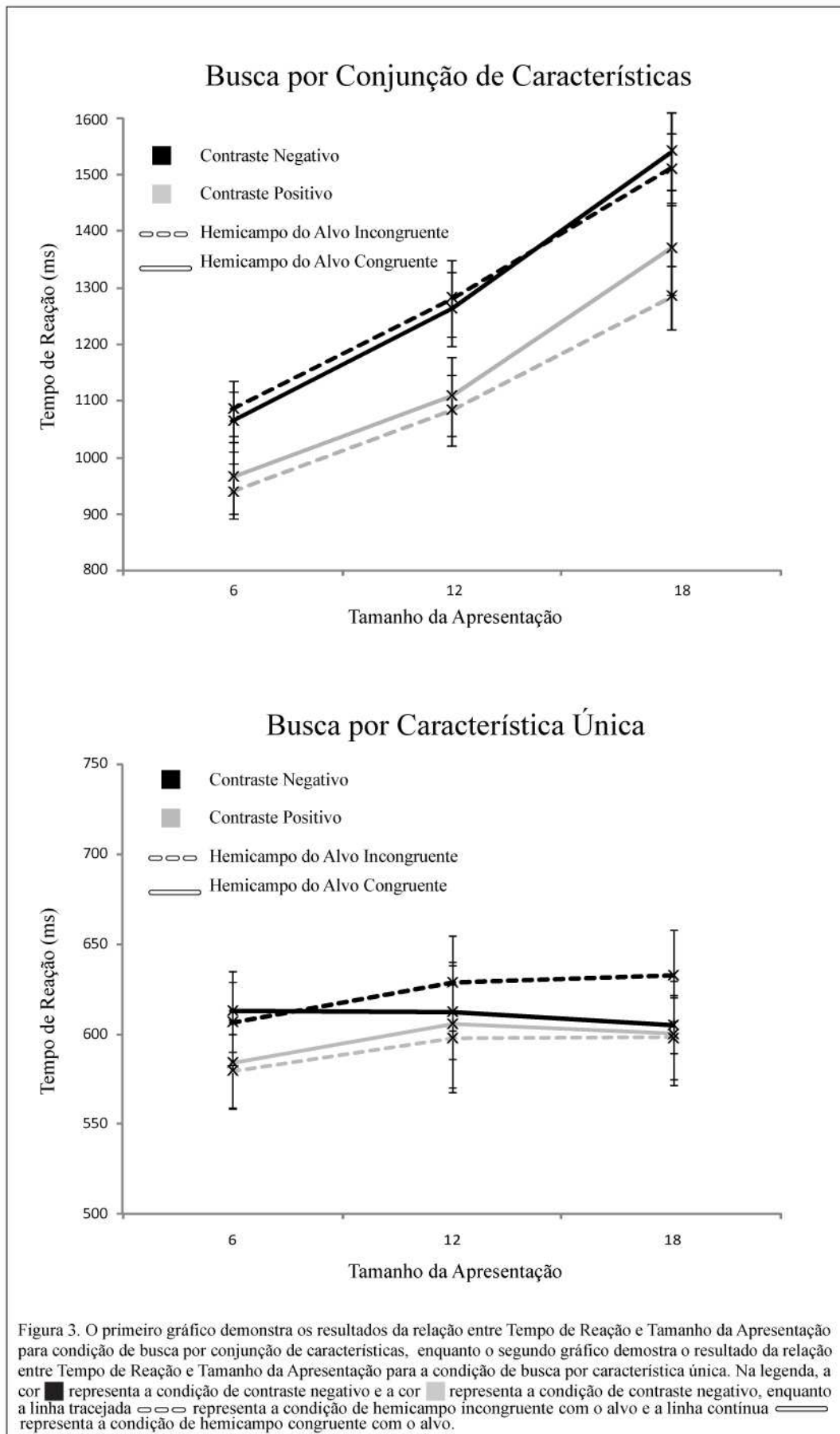
2. Números de elementos relevantes: Composto por três níveis de elementos relevantes (6,12,18);

3. Hemicampo do alvo: Composta por dois níveis, congruente e incongruente com a orientação da textura, ou seja, o alvo podia ser apresentado no mesmo hemicampo para quais as setas apontavam ou no hemicampo contrário.

O tempo de reação e taxa de erros foram registrados através do aplicativo E-prime 2.0 (Psychology Software Tools, Inc).

3.1.3. Resultados e Discussão

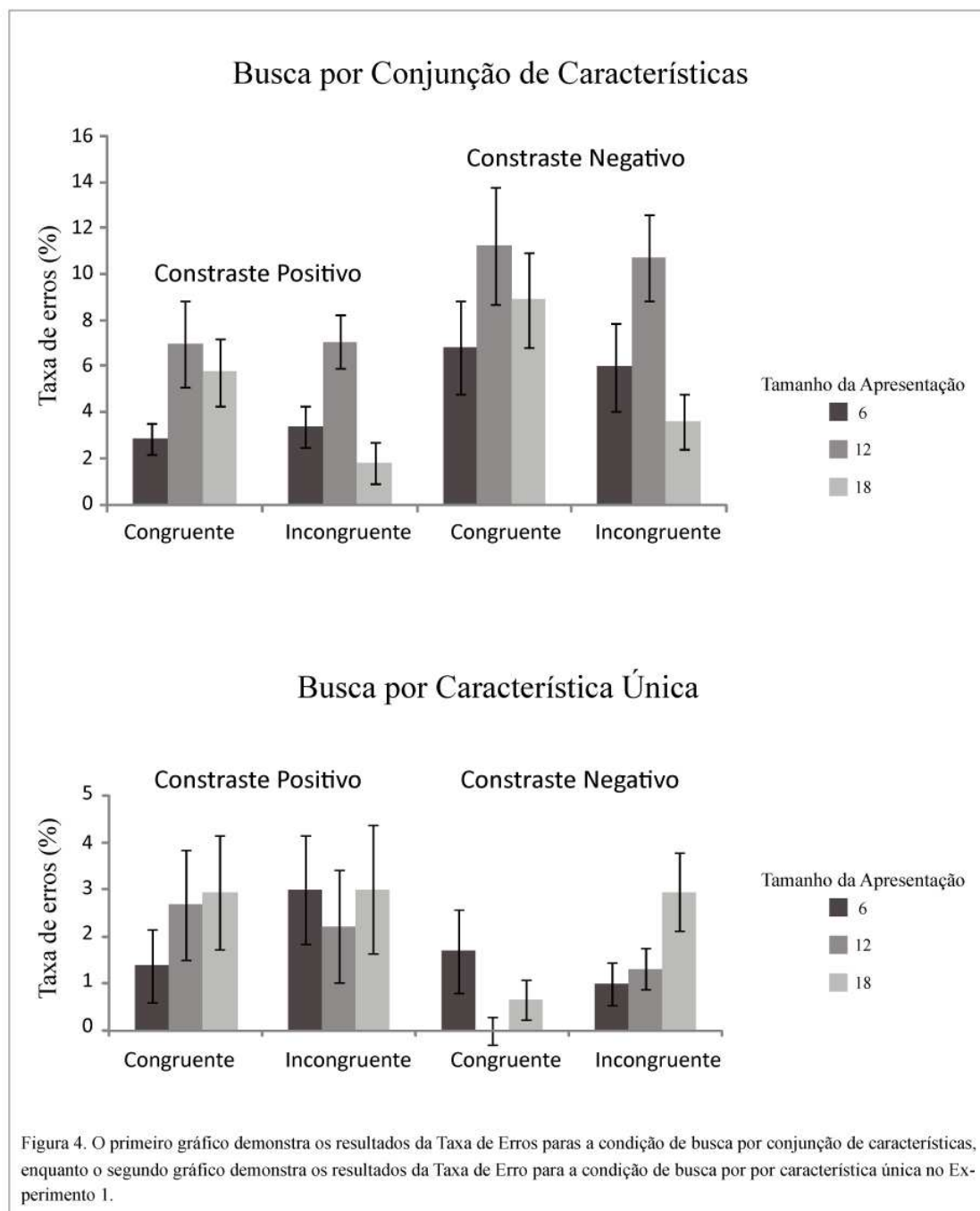
As medidas de tempo de reação que ultrapassaram o valor de dois desvios padrão somado a média foram excluídas da análise. Esse critério de corte foi estabelecido usando as medidas individuais de desvio padrão e média de cada participante, criando um valor de corte específico para cada um. Os dados de tempo de reação e taxa de erro foram então submetidos à análise de variância para medidas repetidas (ANOVA). Os resultados são demonstrados pelos Figura 3 e Figura 4, os gráficos da condição de busca por a características não foram expostos por não apresentarem correlações significativas em geral .



Os resultados mostraram que o tamanho da apresentação influenciou o tempo de

reação na condição de busca por conjunção de características, tanto na condição de contraste positivo ($F(2,14) = 54,60$; $p < 0,001$; $\eta_p^2 = 0,88$), quanto na condição de contraste negativo ($F(2,14) = 91,07$; $p < 0,001$; $\eta_p^2 = 0,92$). Mostrando que o tempo de reação tende a aumentar quando com o aumento no número de estímulos relevantes na busca por conjunção de características. Como esperado, o tamanho da apresentação não influenciou o tempo de reação na condição de busca por característica única, tanto na condição de contraste positivo ($F(2,14) = 1,86$; $p = 0,19$), quanto na condição de contraste negativo ($F(2,14) = 1,09$; $p = 0,36$). Desta maneira, se observa que aumento do número de estímulos relevantes não altera o tempo de reação. O tamanho da apresentação também influenciou a taxa de erros na condição de busca por conjunção de características, tanto na condição de contraste positivo ($F(2,14) = 6,23$; $p < 0,012$; $\eta_p^2 = 0,47$), quanto na condição de contraste negativo ($F(2,14) = 6,18$; $p < 0,012$; $\eta_p^2 = 0,46$). Todavia, não houve influência do tamanho da apresentação na taxa de erros na condição de busca por característica única, tanto na condição de contraste positivo ($F(2,14) = 0,92$; $p = 0,42$), quanto na condição de contraste negativo ($F(2,14) = 3,12$; $p = 0,11$). Esse resultado corrobora a Teoria de Integração de Características, em que a busca por uma característica única ocasiona a uma busca pré-atentiva eficiente sem a influência do tamanho da apresentação, enquanto a busca por uma conjunção de característica gera uma busca atenta, serial e mais lenta influenciada pelo tamanho da apresentação.

Um dado interessante é que a taxa de erros foi maior quando o tamanho da apresentação possuía 18 itens do que quando possuía 12. Uma possível interpretação para esse resultado é uma mudança de estratégia dos participantes para a condição de 18 itens, talvez inferindo a dificuldade da tarefa pelo aumento de itens e antecipando a possibilidade de se equivocarem, os participantes se engajaram em uma estratégia que priorize a precisão em relação ao tempo de execução. Antecipando a dificuldade da



tarefa, os participantes seriam mais “cuidadosos”, gerando um aumento no limiar de ativação da resposta, o que diminuiria os erros e aumentaria o tempo de reação. O participante não utilizaria essa estratégia mais “cuidadosa” na condição de 12 itens, por uma expectativa menor de dificuldade da tarefa.

Quando a condição de busca era por conjunção de características, os participantes tendiam a ser significativamente mais rápidos ($F(1,7) = 53,69$; $p < 0,001$; $\eta_p^2 = 0,88$) e cometerem menos erros ($F(1,7) = 14,96$; $p < 0,006$; $\eta_p^2 = 0,68$) na condição de contraste positiva em relação a de contraste negativo. Quando a busca era por característica única, tendiam a ser significativamente mais rápidos na condição de contraste positivo ($F(1,7) = 6,57$; $p < 0,037$; $\eta_p^2 = 0,48$), porém apenas apresentou uma não apresentou uma influência significativa sobre a taxa de erros ($F(1,7) = 3,796$; $p = 0,092$). Isso acontece possivelmente, pois na condição contraste positivo os estímulos relevantes possuem uma luminância maior do que a textura, facilitando a seleção dos primeiros e o descarte da segunda.

Na busca por conjunção de característica, a média do tempo de reação na condição de hemicampo do alvo congruente com a textura não diferiu significativamente da condição de hemicampo do alvo incongruente com a textura, tanto na condição de contraste positivo ($F(2,14) = 1,22$; $p = 0,32$), quanto na condição de contraste negativo ($F(2,14) = 0,14$; $p = 0,71$). A taxa de erros também não diferiu significativamente entre a condição de hemicampo do alvo congruente e hemicampo do alvo incongruente, tanto na condição de contraste positivo ($F(2,14) = 1,52$; $p = 0,26$), quanto na condição de contraste negativo ($F(2,14) = 2,07$; $p = 0,19$).

O mesmo ocorreu na busca por característica única, a média do tempo de reação na condição de hemicampo do alvo congruente com a textura não diferiu significativamente da condição de hemicampo do alvo incongruente com a textura, tanto

na condição de contraste positivo ($F(2,14) = 0,30$; $p = 0,59$), quanto na condição de contraste negativo ($F(2,14) = 2,30$; $p = 0,17$). A taxa de erros também não diferiu significativamente entre a condição de hemicampo do alvo congruente e hemicampo do alvo incongruente, tanto na condição de contraste positivo ($F(2,14) = 1,28$; $p = 0,29$), quanto na condição de contraste negativo ($F(2,14) = 1,5$; $p = 0,25$).

Os resultados demonstraram que a informação semântica relacionada à textura não teve influência sobre a busca. Esta ausência de influência da informação semântica contrasta com as outras pesquisas usando estímulos pseudo-realistas, em que a informação semântica foi eficiente em orientar a busca. Para explicar essa discrepância foram geradas três hipóteses explicativas.

A primeira se refere à Teoria de Esquema Geral da Cena. Essa teoria propõe que o efeito da informação da informação semântica é decorrente da ativação de um esquema geral nos momentos iniciais da visualização da cena. Este esquema se constitui de representação geral das relações entre objetos da cena que guia a atenção para as áreas semanticamente relevantes. A ativação do esquema geral seria uma propriedade única dos estímulos mais realistas apresentados em um contexto significativo. Já objetos sem relação contextual, como as setas utilizadas, necessitariam que a informação de baixo-nível fosse integrada para ter acesso à informação semântica. Assim, a informação semântica só estaria disponível tardiamente, não exercendo influência sobre a busca. Isto indica um processamento serial da informação semântica na busca visual, quando o estímulo não pertence a um contexto semântico.

A segunda hipótese é da seleção discreta da textura. Apesar da presença da informação semântica, o fato dela ser apresentado em forma de textura (elementos repetidos e em padrão de luminância diferente dos estímulos relevantes) gera sua segmentação em um processo pré-atentivo, resultado já visto em outros estudos sobre

efeito da textura. Como as setas apresentam um padrão de luminância diferenciada dos elementos relevantes, as dicas endógenas não gerariam ativação suficiente para seleção da informação, sendo eficientemente descartadas.

A última hipótese está relacionada à estratégia de busca. É possível que a informação da textura não fosse descartada, e chegasse a ser processada, porém como apenas 50% dos alvos se localizavam em posições congruentes com a informação da textura, a informação semântica não seria informativa. Em outras palavras, o possível direcionamento da atenção gerada pela textura semântica teria a mesma probabilidade de facilitar ou de prejudicar a busca, não sendo, assim, usadas. Já em ambientes naturais, a informação semântica é altamente informativa, tendo uma grande probabilidade de indicar a localização correta do alvo. Assim, o processamento da informação já estaria “configurado” para processar a informação semântica de ambientes naturais. De tal forma, que mesmo quando o desenho experimental com imagens pseudo-realistas usa um controle com a mesma quantidade de provas congruentes e incongruentes, a experiência anterior predispõe ao uso da informação semântica para guiar a busca. Para que a textura fosse informativa, seria necessário que houvesse uma proporção maior de provas com a condição de posição do alvo congruente do que incongruente.

3.2 Experimento 2

Para testar a hipótese que ausência da influência da textura foi devido ao direcionamento das setas não ser informativo, foi delineado o experimento 2. Neste, com hemicampo do alvo é congruente com o direcionamento da textura em 80% das provas, nos outros 20%, o hemicampo do alvo é incongruente. Com a ordem das provas aleatória, os participantes não são capazes de perceber que o número de provas na condição de hemicampo do alvo congruente é maior que o incongruente. Porém, se a

informação semântica está sendo processada, se espera que o haja uma adaptação e a informação semântica passe a influenciar na busca com o avanço das provas.

3.2.1 Participantes

O experimento 2 foi realizado por 10 participantes sendo 4 homens e 6 mulheres, sendo somente um canhoto e apresentando todos visão normal ou corrigida com idade média de 26 anos ($DP= 3,3$). Somente dois participantes do grupo haviam realizado do experimento 1. Novamente, os participantes foram esclarecidos acerca das características do experimento, bem como sobre a preservação do sigilo da sua identidade. Após todos os esclarecimentos necessários, e imediatamente antes do início da tarefa experimental, todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido.

3.2.2 Estímulos

Os estímulos foram, novamente, apresentados em um monitor de 17 polegadas de tubo catódico, configurado em uma resolução de 800x600 pixels com 24 bit de profundidade de cor através do uso do aplicativo E-prime 2.0.

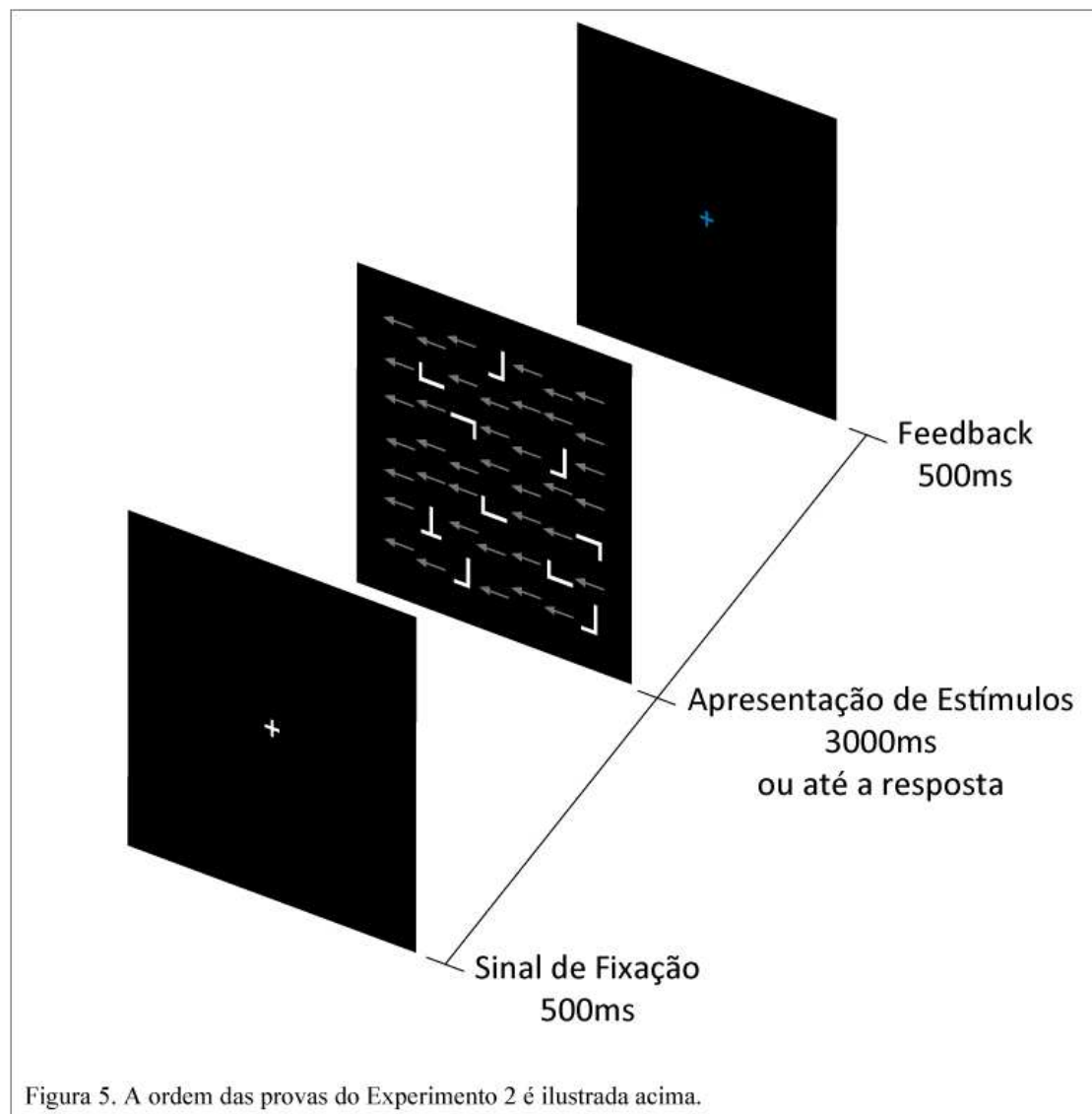
Neste experimento a condição de busca por característica única foi retirada do desenho experimental, por estar claro pelo Experimento 1 que este tipo de busca ocorre pré-atentivamente não havendo tempo hábil para que a informação da textura seja processada. Assim, só foi apresentada a condição de busca por **conjunção de característica**. Retirada esta exceção, todos outros estímulos foram os mesmos utilizados no Experimento 1. Assim o estímulo alvo consistia em um “T” (1,5° de largura e 2° de altura) normalmente orientado em metade das provas e um “T” (1,5° de largura e 2° de altura) rotacionado em 180° na outra metade e os estímulos distratores consistiam em “L”s (1,5° de largura e 2° de altura) rotacionados em 0°, 90°, 180° e 270°

e espelhados horizontal e verticalmente. Foi, novamente, usado em imagens de setas paralelamente orientadas e apontando para um mesmo lado como textura. Os estímulos da busca, como no experimento anterior, eram sempre apresentados no centro da tela em uma matriz com 7 linhas e 7 colunas (com lados medindo 16°), compondo 49 áreas, que quando não estavam ocupadas por um estímulo relevante, eram preenchidas pela textura.

Outra vez, os estímulos apresentavam duas condições de cor e luminância, uma nomeada de **contraste positivo** e a outra de **contrate negativo**. Na condição de contraste positivo, os estímulos relevantes (alvo e distratores) eram apresentados na cor branca, enquanto a textura era apresentada na cor cinza . Na condição de contraste negativo, os estímulos relevantes (alvo e distratores) eram apresentados na cor cinza, enquanto a textura era apresentada em cor branca. Em todas as condições o fundo da cena era apresentado em negro. Os estímulos apresentados na cor branca possuíam 6,9 cd/m² de luminância, os apresentados na cor cinza possuíam 4,1 cd/m² de luminância, enquanto o fundo negro possuía 1,85 cd/m² luminância. O contraste entre os estímulos relevantes e a textura foi de aproximadamente 25%, calculado através da fórmula de Michelson ($C = L_r - L_t / L_r + L_t$, onde: C =contraste da luminância entre os estímulos, L_r = luminância dos estímulos relevantes e L_t =luminância dos elementos da textura).

3.2.3 Procedimentos

Assim como no Experimento 1 os participantes se acomodaram em uma cadeira confortável à uma distância de 57 cm do monitor, em uma sala com luminosidade controlada. A tarefa dos participantes, novamente, era procurar por um determinado alvo e responder, o mais preciso e rápido possível, pressionando a tecla correspondente em um teclado numérico, sendo “1” para “T” em orientação normal e “2” para “T” rotacionado em 180°.



Cada participante realizou duas sessões experimentais em dias distintos. Cada sessão experimental era composta por um bloco experimental de 660 provas. Na primeira sessão, foram apresentados os estímulos na condição de contraste negativo, enquanto na segunda sessão, os estímulos foram apresentados na condição de contraste positivo.

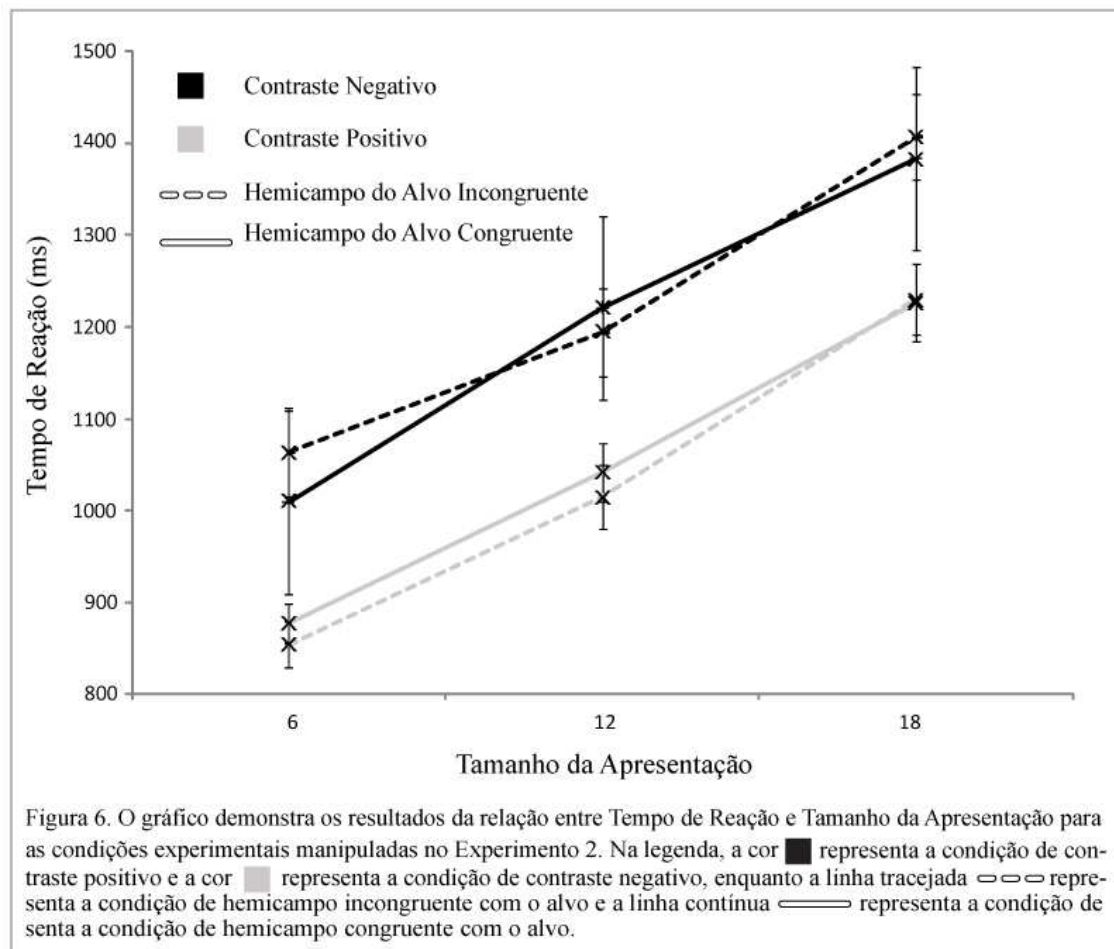
Assim como no Experimento 1, no início da sessão eram apresentadas instruções na tela, orientando o participante acerca da natureza da tarefa. Imediatamente após a leitura das instruções e possíveis esclarecimentos, os participantes realizaram um bloco de treino de vinte provas, nas quais as respostas foram descartadas. Cada prova era

iniciada com a apresentação de um sinal de fixação ocular “+” apresentado no centro da tela por 500ms. Após esse intervalo, o sinal de fixação era substituído pela apresentação dos estímulos de busca que permanecia na tela por 3s ou até que o participante respondia. Caso a resposta fosse correta, era apresentado como feedback o sinal de fixação “+” no centro da tela na cor azul, caso fosse incorreta era apresentado na cor vermelha. Quando a resposta tardava por mais que 3s, um sinal de interrogação na cor verde era apresentado no centro da tela. Esses sinais de feedback perduram por 500ms e a prova era encerrada.

Foram manipulados 3 fatores:

- 1. Números de elementos relevantes:** Composto por três níveis igualmente distribuídos de elementos relevantes (6,12,18);
- 2. Textura composta por setas:** Composta por dois níveis igualmente distribuídos; setas direcionadas para esquerda e setas direcionadas para direita;
- 3. Hemicampo do alvo:** Composta por dois níveis, congruente e incongruente com a orientação da textura, ou seja, o alvo podia ser apresentado no mesmo hemicampo para quais as setas apontavam ou no hemicampo contrário. Em 80% das provas o hemicampo do alvo era congruente com o direcionamento da textura e em 20 % era incongruente.

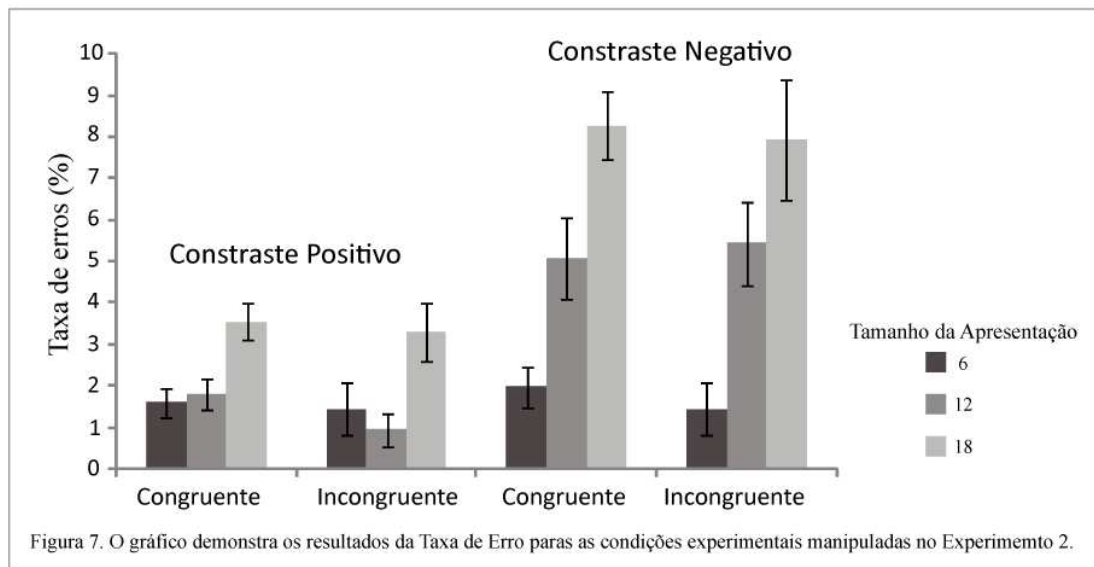
O tempo de reação e taxa de erros foram registrados através do aplicativo E-prime 2.0 (Psychology Software Tools, Inc).



3.2.4 Resultados e Discussão

Novamente, as medidas de tempo de reação que ultrapassaram o valor de dois desvios padrão somado a média foram excluídas da análise. Esse critério de corte foi estabelecido usando as medidas individuais de desvio padrão e média de cada participante, criando um valor de corte específico para cada um. Os dados de tempo de reação e taxa de erro foram então submetidos a análise de variância para medidas repetidas (ANOVA). Os resultados são demonstrados pelas Figura 6 e Figura 7.

Os resultados mostraram que o tamanho da apresentação influenciou o tempo de reação na condição de busca por conjunção de características, tanto na condição de contraste positivo ($F(2,18)=169$; $p<0,001$; $\eta_p^2=0,95$), quanto na condição de contraste negativo ($F(2,18)=217$; $p<0,001$; $\eta_p^2=0,96$). Demonstrando que o tempo de reação tende a aumentar quando com o aumento no número de estímulos relevantes na busca. O



tamanho da apresentação também influenciou a taxa de erros na condição de busca por conjunção de características, tanto na condição de contraste positivo ($F(2,18)=11,89$; $p<0,001$; $\eta_p^2=0,56$), quanto na condição de contraste negativo ($F(2,18)=20,14$; $p<0,001$; $\eta_p^2=0,69$).

Nesse experimento, a taxa de erros aumentou com o acréscimo de itens em todas as condições de tamanho da apresentação, diferente da condição de busca por conjunção de características do experimento 1, em que a condição com 18 itens apresentou uma menor taxa de erros do que a com 12 itens. Isso, a princípio, é contraditório, pois a única diferença do experimento 2 para a condição de busca por conjunção de características do experimento 1 é a porcentagem de provas com a textura congruente com a localização do alvo, o que não explicaria a diferença nos dados. Porém a interpretação apresentada que a diferença na estratégia decorrente em expectativa da tarefa pode acomodar o resultado. No experimento, existia também a tarefa de busca por característica única, sensivelmente mais fácil que a tarefa de busca por conjunção de características. Assim, no experimento 1 a expectativa da dificuldade foi influenciada pela execução dessa tarefa de busca por característica única, o que gerou uma expectativa de dificuldade maior do que a inferida pelos participantes do experimento 2

que não realizaram a tarefa de busca por característica única. Essa diferença no critério de chão para inferência da dificuldade, poderia ter levado aos participantes serem mais “cautelosos” no experimento 1 na condição de 18 itens, do que a mesma condição no Experimento 2.

Novamente, os participantes tendiam a ser mais rápidos ($F(1,9) = 57,79$; $p < 0,001$; $\eta_p^2 = 0,86$), e cometerem menos erros ($F(1,9) = 28,49$; $p < 0,001$; $\eta_p^2 = 0,76$), na condição de contraste positiva em relação a de contraste negativo.

A média do tempo de reação na condição de hemicampo do alvo congruente com a textura não diferiu significativamente da condição de hemicampo do alvo incongruente com a textura, tanto na condição de contraste positivo ($F(2,18) = 1,37$; $p = 0,27$), quanto na condição de contraste negativo ($F(2,18) = 3,29$; $p = 0,10$). A taxa de erros também não diferiu significativamente entre a condição de hemicampo do alvo congruente e hemicampo do alvo incongruente, tanto na condição de contraste positivo ($F(2,18) = 0,07$; $p = 0,41$) quanto na condição de contraste negativo ($F(2,18) = 0,07$; $p = 0,79$). Apesar de em 80% das provas o hemicampo do alvo ser coerente com a textura, os resultados indicam que informação semântica não facilitou a busca. Isto sugere uma seleção eficaz dos estímulos relevantes e segmentação textura baseado em estímulos físicos, não tendo a informação semântica nenhuma influência sobre a seleção dos estímulos, mesmo quando está pode beneficiar a busca.

3.3 Experimento 3

O experimento 3 foi delineado com objetivo de investigar melhor a hipótese de descarte discreta da textura. O seu desenho experimental foi desenvolvido baseado nos estudos sobre pré-ativação (priming) de características na busca visual (Maljkovic & Nakayama, 1994; Maljkovic & Nakayama, 1996; Kristjánsson, Wang e Nakayama, 2002). Nestes estudos, foi observado que quando as características do alvo são comuns

as do alvo de uma busca anterior, isto acelera a busca. Porém, quando as características do alvo são compartilhadas com distratores de uma busca anterior, isto retarda a busca. Assim, a ativação gerada pelas características de um alvo, ou a inibição dos distratores, se estende pelo tempo, influenciando nas buscas das provas seguintes até que essa ativação, ou inibição, se enfraqueça.

Como a ativação ou inibição de um estímulo em uma prova pode ser influenciada pela pré-ativação de características comuns em estímulos de provas anteriores, é possível pensar em um desenho experimental em que o descarte da textura seja prejudicado, ao mesmo tempo em que é aumentada a probabilidade dessa ser selecionada pela atenção. Isso é possível, através da apresentação de uma primeira prova em que as características do alvo fossem compartilhadas com os elementos da textura semântica da uma segunda prova que a seguisse.

Desta maneira, delineou-se um experimento com provas compostas por duas tarefas de busca. Em uma primeira busca, os estímulos relevantes são compostos por setas, sendo o alvo uma seta com triângulo deslocado e os distratores outras setas normais, enquanto a textura é composta “L”s rotacionados em diversas orientações. Na segunda busca, as setas são parte da textura como em experimentos anteriores, enquanto o alvo é um “T” e os distratores são “L”s. Espera-se com este delineamento que as características que compõem a seta sejam ativadas pela primeira busca e essa ativação se estenda para segunda busca.

Como estratégia de controle foi criada uma segunda condição de pares de busca. Nesta, em ambas as buscas os estímulos relevantes são o alvo “T” e “L”s como distratores e a textura é composta por setas. O que diferencia as buscas é que na primeira as setas possuem uma haste e duas pontas, assim, não há um hemisfério específico em que as setas estejam direcionadas. Enquanto na segunda, as setas possuem

uma haste e somente uma ponta direcionada para um hemisfério específico. A proporção de provas com pares sem ativação das setas foi de 60%, enquanto a de pares com ativação da seta foi de 40%. Isso foi feito para controlar a influência da ativação das características das setas.

A hipótese inicial é que as provas com pares com ativação de setas teriam sua segunda busca mais lenta do que a segunda busca de pares sem ativação de seta, devido ao processo de seleção estar ainda sob influência da primeira busca. Como as características que compõem as setas estão ativadas pela busca anterior, aumenta a probabilidade dessas serem selecionadas pela atenção e a informação semântica guiar a busca.

3.3.1 Participantes

O experimento 3 foi realizado por 15 participantes sendo 6 homens e 9 mulheres, todos destros e apresentando visão normal ou corrigida, com idade média de 26 anos ($DP = 3,7$). Os participantes foram esclarecidos acerca das características do experimento, assim como sobre a preservação do sigilo da sua identidade. Após todos os esclarecimentos necessários todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido.

3.3.2 Estímulos

Os estímulos foram, novamente, apresentados através do aplicativo E-prime 2.0 em um monitor de 17 polegadas de tubo catódico, configurado em uma resolução de 800x600 pixels com 24 bit de profundidade de cor.

Neste experimento usaram-se como estímulos “T” (1,5° de largura e 2° de altura) e “L” (1,5° de largura e 2° de altura) e setas (com haste de 2,3° e ponta com 0,8° de altura) como nos experimentos anteriores. Os estímulos da busca eram sempre

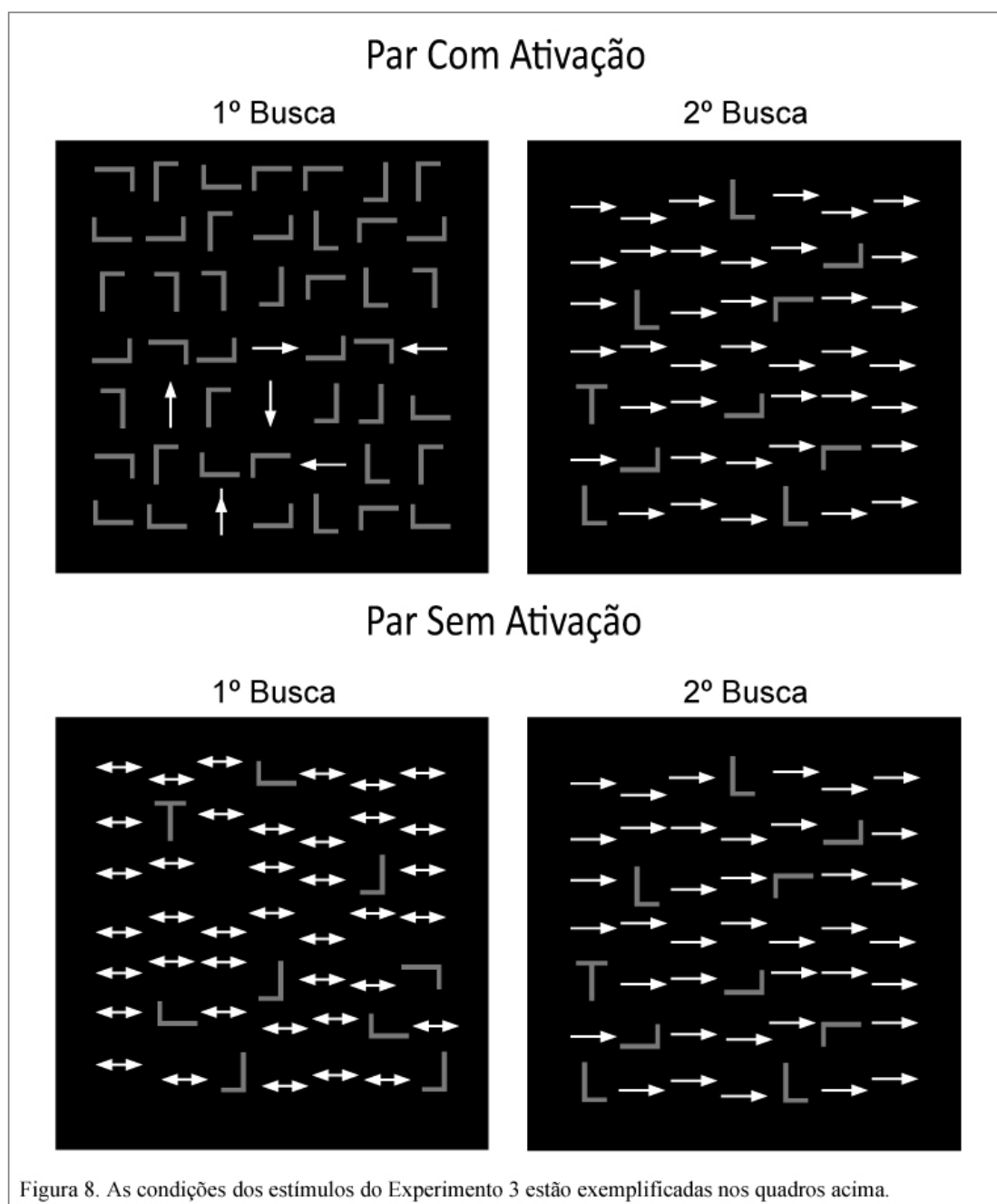
apresentados no centro da tela em uma matriz imaginária com 7 linhas e 7 colunas (com 16 ° de lado) compondo 49 áreas, que quando não estavam ocupadas por um estímulo relevante, eram preenchidas pela textura.

Foram estabelecidas duas condições de prova, sendo cada uma composta por um par de tarefas de busca. Em ambos os pares, a segunda tarefa de busca era a mesma, no caso, uma busca por “T”. A diferença entre as condições de prova residia na primeira tarefa de busca. Esta primeira tarefa, em uma condição, era uma busca cujo o alvo era uma seta com triângulo deslocado, e em outra condição, era uma busca por “T”. Assim, as condições foram nomeadas de **par com ativação das setas** e **par sem ativação de setas**, abaixo estão descritos os estímulos apresentados em cada uma dessas condições:

1. Par com ativação de setas: Na primeira tarefa de busca, o estímulo alvo consistia em uma seta com o triângulo descolado 0,3° graus presente em metade das provas e ausente nas demais. Os distratores eram setas rotacionadas em 0°, 90°, 180° e 270°. A textura era composta por “L”s rotacionados em 0°, 90°, 180° e 270° e espelhados horizontal e verticalmente. Na segunda tarefa de busca, o estímulo alvo o estímulo alvo consistia em um “T” normalmente orientado em metade das provas e um “T” rotacionado em 180° na outra metade, os estímulos distratores consistiam em “L”s rotacionados em 0°, 90°, 180° e 270° e espelhados horizontal e verticalmente. A textura usada consistia em imagens de setas paralelamente orientadas e apontando para um mesmo lado.

2. Par sem ativação de setas: Na primeira tarefa de busca o estímulo alvo o estímulo alvo consistia em um “T” normalmente orientado em metade das provas e um “T” rotacionado em 180° na outra metade. Os estímulos

distratores consistiam em “L”s rotacionados em 0°, 90°, 180° e 270° e espelhados horizontal e verticalmente. A textura usada consistia em imagens de setas com uma haste e duas pontas paralelamente orientadas. Na segunda tarefa de busca, o estímulo alvo o estímulo alvo consistia em um “T” normalmente orientado em metade das provas e um “T” rotacionado em 180° na outra metade. Os estímulos distratores consistiam em “L”s rotacionados



em 0°, 90°, 180° e 270° e espelhados horizontal e verticalmente. As texturas usadas consistiam em imagens de setas paralelamente orientadas e apontando para um mesmo lado.

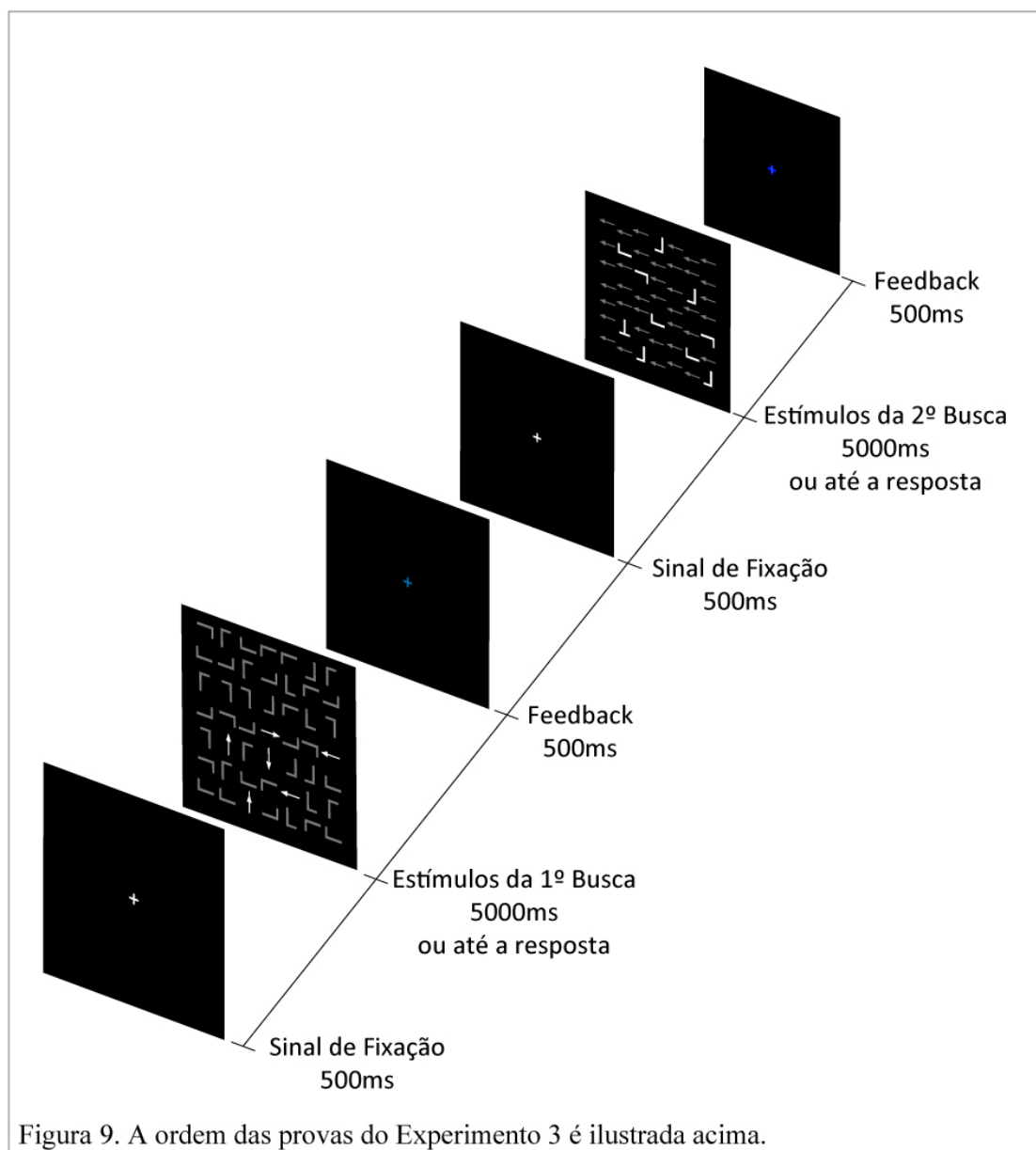
Os estímulos relevantes eram apresentados na cor branca e a textura era apresentado na cor cinza sobre o fundo branco em todas as condições. Os estímulos apresentados na cor branca possuíam 6,9 cd/m² de luminância, os apresentados na cor cinza possuíam 4,1 cd/m² de luminância, enquanto o fundo negro possuía 1,85 cd/m² luminância. O contraste entre os estímulos relevantes e a textura foi de aproximadamente 25%, calculado através da fórmula de Michelson ($C = \frac{L_r - L_t}{L_r + L_t}$, onde: C =contraste da luminância entre os estímulos, L_r = luminância dos estímulos relevantes e L_t =luminância dos elementos da textura). A Figura 8 apresenta um exemplo esquemático dos estímulos utilizados.

A Figura 8 apresenta um exemplo esquemático dos estímulos utilizados.

3.3.3 Procedimentos

Os participantes foram acomodados em uma cadeira confortável a uma distância de 57 cm do monitor em uma sala com luminosidade controlada. A tarefa dos participantes, novamente, era procurar por um determinado alvo e responder, o mais preciso e rápido possível, pressionando a tecla correspondente em um teclado numérico. As teclas a serem pressionadas variavam de acordo com a condição de tarefa de busca efetuada, sendo na condição de busca por “T”, “1” para “T” em orientação normal e “2” para “T” rotacionado em 180°, enquanto na condição de busca por setas o participante deveria pressionar “1” quando a seta com triângulo deslocado estivesse presente e “2” para quando estivesse ausente.

No início da sessão eram apresentadas instruções na tela com orientações sobre a tarefa. Imediatamente após a leitura das instruções e possíveis esclarecimentos, os participantes realizaram um bloco de treino de onze provas, em que as respostas foram descartadas. Como a tarefa deste experimento era mais complexa do que as tarefas exigidas nos dois primeiros experimentos, eram novamente esclarecidas às possíveis dúvidas do participante, caso fosse constatado que este não compreendia as instruções, novas instruções e esclarecimentos eram feitos e um novo bloco de treino com onze provas era realizado, em que as respostas foram novamente descartadas. Em seguida o



participante respondia a dois blocos experimentais, cada um contendo 200 provas, separados por um intervalo de 10 minutos.

Cada prova era iniciada com a apresentação de um sinal de fixação ocular “+” no centro da tela por 500ms. Após esse intervalo, o sinal de fixação era substituído pela apresentação dos estímulos da primeira tarefa de busca que permanecia na tela por 5s ou até que o participante respondia. Caso a resposta fosse correta, era apresentado como feedback o sinal de fixação “+” no centro da tela na cor azul, caso fosse incorreta era apresentado na cor era vermelha. Quando a resposta tardava por mais que 5s, um sinal de interrogação na cor verde era apresentado no centro da tela. Esses sinais de feedback perduram por 500ms. Depois o sinal de feedback era substituído pelo sinal de fixação ocular “+” no centro da tela, durando por 500ms. Após esse intervalo, o sinal de fixação era substituído pela apresentação dos estímulos da primeira tarefa de busca que permanecia na tela por 5s ou até que o participante respondia. Novamente, caso a resposta fosse correta, era apresentado como feedback o sinal de fixação “+” no centro da tela na cor azul, caso fosse incorreta era apresentado na cor era vermelha. Quando a resposta tardava por mais que 5s, um sinal de interrogação na cor verde era apresentado no centro da tela. Esses sinais de feedback perduravam por 500ms e a prova era encerrada. A Figura 9 ilustra a ordem da prova.

Foram manipulados 4 fatores:

- 1. Números de elementos relevantes:** Composto por dois níveis igualmente distribuídos de elementos relevantes (6,12);
- 2. Primeira busca do par:** Composto por dois níveis. Busca por seta com triângulo deslocado, que pode estar presente ou ausente e busca por “T”, que pode estar na orientação normal ou deslocado. A proporção entre esses níveis foi de 40% das provas com busca por setas e 60% com busca por “T”.

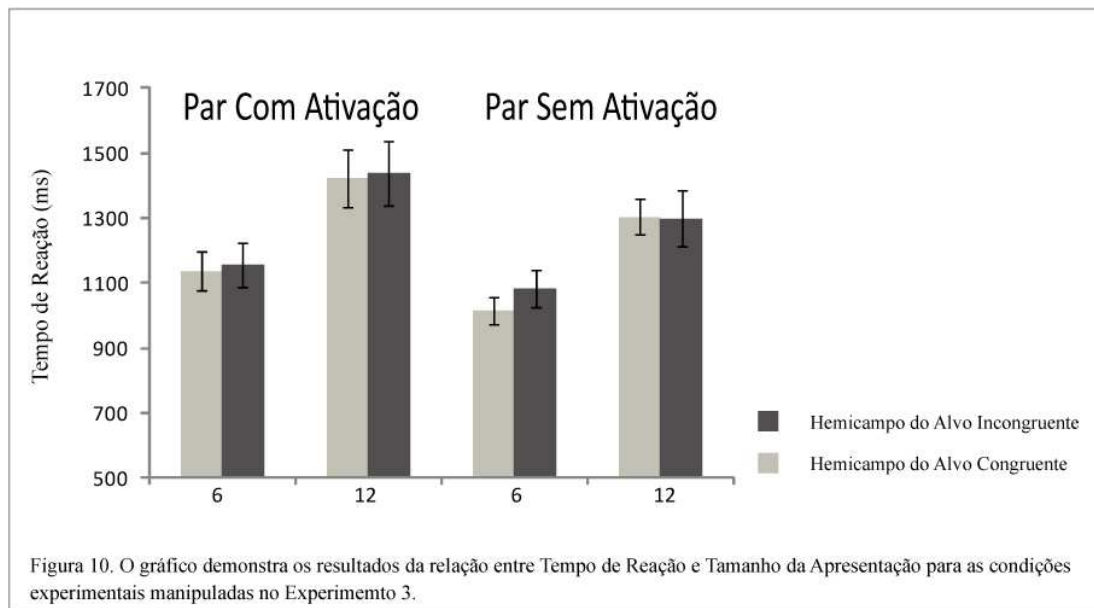
3. Textura composta por setas: Implementado somente na segunda tarefa de busca do par, composta por dois níveis igualmente distribuídos; setas direcionadas para esquerda e setas direcionadas para direita;

4. Hemicampo do alvo: Implementado somente na segunda tarefa de busca do par, composta por dois níveis igualmente distribuídos, congruente e incongruente com a orientação da textura, ou seja, o alvo podia ser apresentado no mesmo hemicampo para quais as setas apontavam ou no hemicampo contrário.

O tempo de reação e taxa de erros foram registrados através do aplicativo E-prime 2.0 (Psychology Software Tools, Inc).

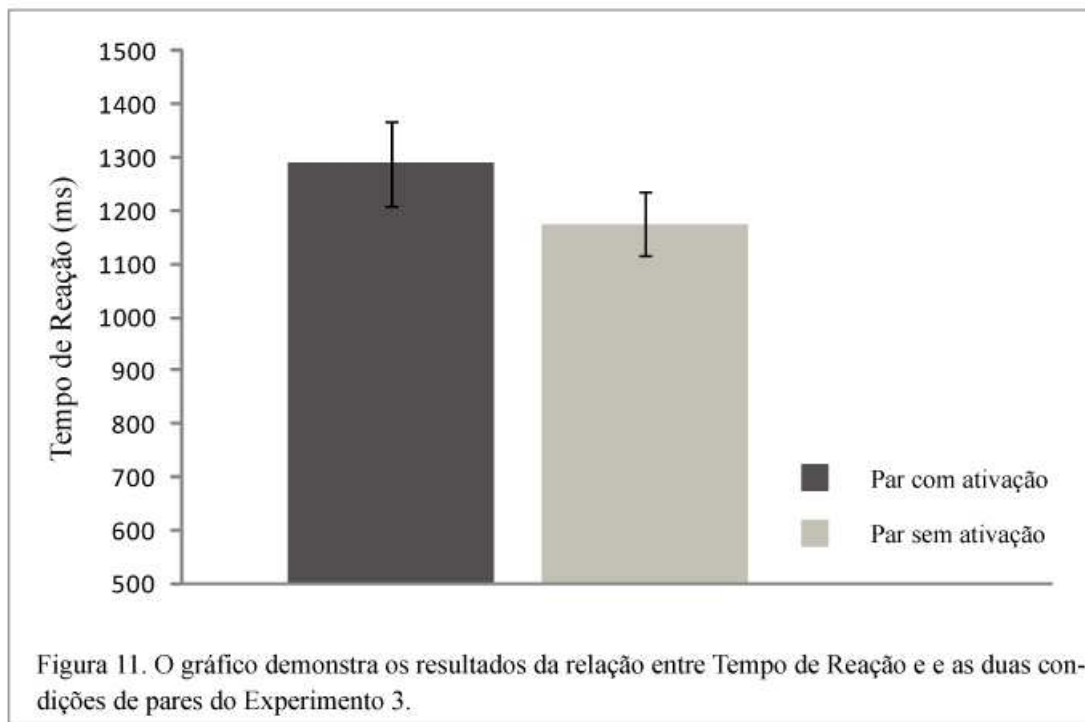
3.3.4 Resultados e Discussão

Novamente, as medidas de tempo de reação que ultrapassaram o valor de dois desvios padrão somado a média foram excluídas da análise. Esse critério de corte foi estabelecido usando as medidas individuais de desvio padrão e média de cada participante, criando um valor de corte específico para cada um. Os dados de tempo de reação e taxa de erro foram então submetidos à análise de variância para medidas repetidas (ANOVA). Os resultados são demonstrados pelas Figura 10 e Figura 11, a variável taxa de erros não foi exposta por falta de correlações significativas em geral.



Os resultados mostraram que o tamanho da apresentação influenciou o tempo de reação ($F(1,14) = 87,58$; $p < 0,001$; $\eta_p^2 = 0,86$). Mostrando que o tempo de reação tende a aumentar quando com o aumento no número de estímulos relevantes. Porém, diferente dos experimentos anteriores, o tamanho da apresentação não exerceu influência sobre a taxa de erros ($F(1,14) = 0,77$; $p = 0,39$).

A média do tempo de reação na condição de par com ativação de setas foi significativamente maior do que na condição de par sem ativação de setas ($F(1,14) = 16,89$; $p < 0,001$; $\eta_p^2 = 0,54$). Isto sugere, que como planejado, as características que compõem as setas foram ativadas pela primeira busca e essa ativação se estendeu para segunda busca, dificultando o descartaste da textura na segunda busca na condição de par com ativação de setas. Porém, na taxa de erros não houve o mesmo padrão, A quantidade de erros não foi significativamente maior em nenhuma das condições ($F(1,14) = 1,21$; $p = 0,28$).



Os resultados demonstraram, por fim, que não houve diferença significativa no tempo de reação entre duas condições de hemicampo do alvo (congruente e incongruente com a textura) tanto na condição de par com ativação das setas ($F(1,14)=0,53$; $p=0,47$), como na condição de par sem ativação das setas ($F(1,14)=0,95$; $p=0,34$). O mesmo foi observada na taxa de erros, que não apresentou diferença significativa entre as duas condições de hemicampo do alvo na condição de par com ativação ($F(1,14)=0,53$; $p=0,47$) ou par sem ativação ($F(1,14)=0,95$; $p=0,34$). Este resultado aponta para que mesmo com o aumento da ativação da textura, esse não produziu um efeito semântico sobre a busca. Indicando que estímulos semânticos não são capazes de produzir o mesmo efeito observado em estímulos semânticos complexos e pseudo-realistas.

4. Discussão Geral

Os estudos iniciais sobre a atenção seletiva apresentam maior ênfase sobre a informação de baixo-nível e médio-nível, investigando como as características físicas dos estímulos influenciam na busca visual (Duncan & Humphrey, 1989; Treisman & Gelade, 1980; Wolfe, Cave, & Franzel, 1989). No entanto, os modelos de busca visual, abordados nesses estudos, dão pouco destaque à informação de alto-nível, prevendo que a busca visual com estímulos realistas possui uma baixa eficiência, devido à complexidade dos estímulos e alta heterogenidade entre estímulos não relevantes. No entanto, estudos sobre a informação visual de alto-nível demonstraram que a busca com estímulos complexos pseudo-realistas são mais rápidas e eficientes do se previa (Biederman, 1972; Biederman, Mezzanotte & Rabinowitz, 1982; Henderson, Weeks & Hollingworth, 1999; Gordon, 2004; Underwood & Foulsham, 2006; Underwood, Templeman, Lamming, e Foulshan, 2008). A razão para isso é que a informação de alto-nível é capaz de guiar a atenção, facilitando a seleção da informação semanticamente relevante.

O objetivo desse estudo era replicar o efeito da informação semântica usando estímulos físicos simples e investigar se a facilitação gerada pela informação de alto-nível é exclusiva de estímulos realistas apresentados em um contexto semântico, ou um efeito geral da informação de alto-nível. Nos experimentos, um conjunto de setas paralelamente orientadas e apontadas em uma mesma direção foi usado como elemento semântico, a hipótese era que se a informação das setas guiasse a atenção, a busca seria mais eficiente quando o direcionamento das setas fosse congruente com o hemisfério da localização do alvo. No Experimento 1, os resultados dos experimentos demonstraram que quando apresentada em forma de textura com luminância distinta dos elementos relevantes, a informação semântica não influencia a busca. Uma das

hipóteses alternativas estava relacionada à informatividade da textura. No experimento 2, a informatividade da textura foi manipulada de forma que em 80% das provas as setas indicavam o hemisfério do alvo. Porém, não houve benefício por essa informação e os resultados mostraram que a informação semântica não influenciou na busca. Por fim, no Experimento 3, a ativação da textura foi manipulada através de uma pré-ativação gerada por uma busca anterior. Os resultados sugerem que a pré-ativação das características das texturas retardaram a busca de uma forma geral, porém não houve diferença significativa entre as condições de congruência da textura e a localização do alvo.

Os estudos sobre o fundo e textura na busca visual sugerem que os elementos de textura são segmentados dos estímulos relevantes em um estágio pré atencional, podendo influenciar o estágio atencional interagindo com o tamanho da apresentação quando ocorre uma segmentação imperfeita, quanto maior a similaridade física entre os elementos da textura e os estímulos relevantes, maior a probabilidade desta agir retardando o processo de seleção e busca. Assim, as características físicas teriam uma forte influência na segmentação dos estímulos relevantes do fundo, sendo a informação semântica só processada quando selecionada por um processo guiado por características físicas (Rossini & Galera, 2008; Wolfe, Oliva, Horowitz, Butcher & Bompas, 2002).

Neste trabalho, como a textura foi apresentada com luminância distinta dos estímulos relevantes, a ausência de efeito semântico pode ser explicada por uma segmentação pré-atencional dos elementos da textura. Devido a esta segmentação inicial, as setas que compõem a textura não são selecionadas pela atenção, assim a informação semântica não chega a ser processada. Isto indica que quando usados estímulos simples a informação de alto-nível é processada tardiamente, e só pode ser acessada após integração da informação de baixo-nível. Um bom indicativo que a seleção dos estímulos é guiada inicialmente pelas características físicas, ocorrendo de maneira pré-

atentiva e sem a influência da informação semântica, é o resultado do Experimento 2. Apesar da textura ser coerente com a localização do alvo em 80% dos casos, podendo auxiliar no objetivo da tarefa de localizar o alvo mais rapidamente, o sistema atento é incapaz de se beneficiar desta informação. Como a informação de alto-nível está condicionada a uma seleção dos estímulos mediada pelas características físicas, mesmo que a informação semântica seja relevante para a tarefa, ela é descartada caso as características físicas do estímulo que abriga a informação semântica não seja relevante para a tarefa.

Por sua vez, o resultado do Experimento 3 demonstra de forma interessante o efeito das características físicas na seleção da informação. No seu desenho experimental, a textura ainda possuía baixa similaridade física com os estímulos relevantes, assim espera-se que esta seja descartada de forma eficiente. Porém, as provas do experimento se constituíam de pares de buscas, sendo que em 40% das provas os estímulos relevantes da segunda busca compartilhavam características com a textura da primeira, enquanto a textura da segunda possuía características similares aos estímulos relevantes da primeira busca. O objetivo era que o padrão de ativação e inibição que influenciou a seleção da primeira busca se estendesse à segunda, gerando assim, uma interferência no processo de descarte da textura. Nos estudos sobre pré-ativação (priming) de características na busca visual, quando o estímulo alvo de uma busca compartilha características físicas com os distratores de uma busca anterior há um aumento nas médias do tempo de reação, no entanto, quando o estímulo alvo possui características físicas similares ao alvo de uma busca anterior ocorre uma redução nas médias do tempo de reação (Kristjánsson, Wangb e Nakayama, 2002; Maljkovic & Nakayama, 1994; Maljkovic & Nakayama, 1996). Esse efeito foi demonstrado para busca por característica única, assim como conjunção de características, porém somente

eram usados estímulos alvos e distratores sobre um fundo limpo. No experimento 3, quando a segunda busca era precedida por uma busca com as características físicas dos estímulos relevantes e textura invertido, a busca era significativamente mais lenta. Já quando a segunda busca era precedida por uma busca com características físicas similares entre os estímulos relevantes, assim como a textura, a busca era significativamente mais rápida. Isto sugere que o efeito de pré-ativação também é válido para elementos de fundo como a textura. Assim, o processo de segmentação da textura e seleção dos estímulos não é algo isolado e estanque, e sim, uma extensão de processos anteriores, sendo influenciado pelo padrão segmentação e seleção de tarefas próximas temporalmente.

Porém, estudos usando estímulos mais realistas e contextualizados demonstram que a informação de alto-nível tem uma influência tão forte quanto a informação de baixo nível na seleção da informação. Um estudo realizado por Gordon (2004) demonstrou que é possível verificar a influência da informação semântica na alocação da atenção em momentos iniciais da busca, a 40 ms e 70 ms após a exposição da cena. Avaliando os potenciais cerebrais em uma tarefa go/nogo(responder/nãoresponder) em que o alvo era a presença de um animal na cena, Rousselet et. al., (2002) encontrou em aproximadamente 150ms uma diferença significativa nos potenciais, indicando que a informação semântica já havia sido processada. Em estudos realizados por Underwood e Foulsham (2006) e Underwood et al. (2008) a influência da informação semântica teve um peso maior no processo de seleção da informação do que a saliência física dos estímulos. Esses experimentos apresentam dados que demonstram que a informação semântica pode influenciar a seleção da informação no estágio pré-atentivo.

A Teoria de Memória Conceitual de Curto-Prazo (Potter, 1993; 1999), prevê que a possibilidade da informação semântica é processada rapidamente, ela não discrimina

entre a característica do estímulo, no entanto, parece existir uma diferença crítica entre como a informação de alto-nível é processada em desenhos experimentais usando estímulos simples e artificiais e desenhos experimentais que usem estímulos mais realistas com uma maior validação ecológica.

Underwood e Foulsham (2006) explicam o que acarreta o peso maior da informação semântica do que a saliência física é o nível de direcionamento cognitivo da tarefa, o conhecimento anterior do alvo a ser localizado que influencia na busca, gerando a ativação das áreas semanticamente relevantes. De tal forma, que em uma tarefa de memorização realizada pelos autores, a saliência física teve maior peso na seleção inicial do que a informação semântica, pois não havia um alvo específico a ser localizado. Nos experimentos realizados a textura composta de setas, além de possuir um padrão de luminância distinto dos estímulos relevantes, não apresentava nenhuma relação semântica com a tarefa principal de localização com a tarefa principal que era localizar uma determinada letra. Assim, nos experimentos realizados neste trabalho, a busca por uma letra como alvo provavelmente não gerava nenhuma ativação nas setas por uma diferença na luminância, mas também por não apresentar nenhuma associação semântica com o alvo a ser localizado.

Outra hipótese explicativa se refere à Teoria de Esquema Geral da Cena (Biederman, 1972; Biederman, et al. 1982). Essa teoria prevê a ativação de uma representação geral das relações entre objetos nos momentos iniciais da cena que guia a atenção para as áreas semanticamente relevantes. Essa representação da cena está ligada à exposição a ambientes estruturalmente redundantes, de tal forma que o padrão de interação entre objetos é apreendido e estruturado na forma de uma esquema geral. Assim, o que guia a busca não é a informação semântica do objeto em si processado após a integração da sua informação de baixo-nível, mas as relações gerais subjacentes

entre os objetos em uma cena ativada no momento inicial da visualização. Como as setas nos experimentos deste estudo são apresentadas fora de um contexto, não há um esquema a ser ativado, assim, e o único acesso a informação semântica é através da integração da informação de baixo-nível, dessa maneira a seleção da informação é unicamente baseada nas características físicas dos estímulos.

Mas apesar da validade dessas explicações e a capacidade delas de interpretar resultados, ainda existe certo conflito. Com o uso de estímulos simples a seleção parece basear-se unicamente em características físicas, porém com estímulos mais realistas a seleção é guiada pelas características semânticas, tendo neste contexto um efeito muito significativo. Porém, nos estímulos mais simples e a informação de alto-nível é processada só tardiamente e tendo um custo temporal muito maior do que se verifica com uso de estímulos realistas, onde a informação de alto-nível é processada com uma velocidade tão surpreendente, alguns experimentos mostram valores entre 50ms-150ms (Gordon, 2004; Rousselet *et al.*, 2002), que parece estar disponível quase imediatamente. Porém, as hipóteses explicativas por partirem do conceito fundamental que os estímulos naturais teriam alguma propriedade especial que facilitaria a sua detecção, algum efeito sinérgico ou ecológico impossível de reproduzir com estímulos artificiais, criam uma ruptura entre o funcionamento da atenção em diferentes contextos, em uma a informação de alto-nível é processada tardiamente em um estágio serial demandando alto custo temporal, e em outro a informação de alto-nível é processada rapidamente em um estágio paralelo com pouco custo temporal.

No entanto, quando observado a partir da hipótese que a informação de baixo nível age como mediadora para a informação de alto-nível (Evans & Treisman, 2005), o porquê desta diferenciação de efeitos da informação semântica entre estímulo simples e estímulos mais complexos se torna mais claro. Certas características físicas de

complexidade moderada seriam associadas a certas categorias semânticas, sendo possível fazer distinções gerais sem a necessidade que os objetos serem completamente identificados. Por exemplo, a forma de uma asa poderia ser usada para detectar rapidamente a categoria ave, sem que todas as características do objeto sejam integradas, e se saiba exatamente qual ave se trata. Os estudos de Oliva e Torralba (2001) demonstram como a cena pode ser inferida através de características físicas globais da imagem. Essa hipótese insere uma continuidade entre o funcionamento da atenção seletiva nos dois contextos. Nos estímulos simples e artificiais não haveria qualquer influência da informação semântica, pois não existe nenhuma característica que possa agir como mediadora, sendo a informação de alto-nível só acessada pela integração das características. Já os estímulos mais realistas, certos elementos genéricos da informação estariam disponíveis pela detecção de características físicas associadas a essa. E, além disso, a hipótese de que elementos de alto-nível podem ser discriminados a partir da detecção de elementos de baixo nível, propicia uma explicação mais concreta, mensurável e heurística. O que gera a facilitação na busca com estímulos naturais não é mais uma propriedade ecológica desconhecida ou de descrição vaga, e sim características físicas que podem ser especificadas e mensuradas. Nos experimentos realizados nesse trabalho, as setas não apresentavam qualquer informação de baixo-nível que possa indicar sem que seja necessária a integração a localização do alvo, por isso a seleção é baseada somente nas características físicas.

Além de possuir pontos congruentes com as teorias de busca visuais construídas baseadas em estudos com estímulos simples, a hipótese que características físicas de complexidade moderada podem ser aprendidas e usadas para discriminar objetos sem a necessidade de uma representação completa e específica se alinha muito bem com a Teoria de Memória Conceitual de Curto-Prazo (Potter, 1993;1999), ou MCCP. Como

visto nos estudos de Evans e Treisman (2005), as características fornecem discriminações genéricas e inespecíficas, de tal forma, que propicia a rápida detecção de animais. Porém, a presença de distratores humanos que compartilham características com a categoria de animais diminui a eficiência na detecção. A representação não é específica o suficiente para gerar uma discriminação eficaz entre humanos e animais. A representação pré-atentiva dos objetos na TIC é geralmente descrita com um conjunto sem forma de características pouco estruturada e transitória. A idéia central da M CCP é que os estímulos semânticos geram uma ativação genérica e inespecífica nas redes semânticas que esvanece rapidamente se não estruturada, porém esta informação é capaz de influenciar o processo de seleção da informação. Então, se a informação semântica descrita pela M CCP é ativada pelas características físicas descritas pela TIC, é natural que esta seja genérica e pouco estruturada, pois assim, são as características físicas. Por serem, gerados através de características físicas simples, a informação de alto-nível não gera uma representação duradora caso não integradas, isto explicaria a fugacidade da informação semântica na M CCP.

Por fim, este estudo teve um caráter fortemente exploratório por usar estímulos simples e artificiais para estudar um fenômeno normalmente relacionado a estímulos mais realistas. De forma geral, os resultados corroboram outros estudos sobre o efeito da textura sobre a busca visual e sugere que a manipulação da informação semântica usando estímulos físicos não influencia na busca visual como estímulos mais realistas e contextualizados.

Apesar das férteis contribuições, o estudo sobre a atenção seletiva e como a informação de baixo-nível media respostas relacionadas à informação de alto-nível, a pesquisa no campo ainda é limitada. Perguntas sobre quais características podem ser

relacionadas à quais categorias e sobre qual é o grau de especificidade dessas relações, ainda esperam para serem respondidas.

Referências Bibliográficas

Biederman, I. (1972). Perceiving Real-World Scenes. *Science* , 177, 77-80.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene Perception: detecting and judging objects undergoing relational violations. *Cognitive Psychology* , 14, 143-177.

Brockmole, J. R., & Henderson, J. m. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition* , 1 (13), 99-108.

Brockmole, J. R., Henderson, J. M., & Castelano, M. S. (2006). Contextual Cueing in Naturalistic Scenes: Global and Local Context. *Journal of Experimental Psychology* , 32 (4), 699-706.

Chun, M., & Jiang, Y. (1998). Contextual cueing: implicit learning and memory of visual context guides spatial. *Cognitive Psychology* , 36, 28–71.

De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effect of scene context on object identification. *Psychological Research*, 52, 317-329.

Duncan, J & Humphrey, G. W. (1989). Visual Search and stimulus similarity. *Psychological Review*, 96, 433-458

Enns, J. & Resnick, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, 98, 335-351

Evans, K. K., & Treisman, A. (2005). Perception of Objects in Natural Scenes: Is it Really Attention Free? *Journal of Experimental Psychology* , 31 (6), 1476-1492.

Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. New York: Wiley.

Henderson, J., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, 50, 243–271.

Gordon, D. R. (2004). Attentional allocation during the perception of scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 30(4), 760-777

Henderson, J., Weeks, P., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 210-228.

Hollingworth, A., & Henderson, J. M. (1998). Does Consistent Scene Context Facilitate Object Perception. *Journal of Experimental Psychology*, 127 (4), 398-415.

Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cueing: Configuration or Individual Locations. *Perception & Psychophysics*, 66, 454-463.

Kristjánsson, A., Wang, D. & Nakayama, K. (2002). The role of priming in conjunctive visual search. *Cognition*, 85, 37–52.

Loftus, G. R. & Mackworth, N. H. Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 565-572

Levin, D. T., Takarae, Y., Miner, A. G., & Keil, F. (2001). Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception & Psychophysics*, 63, 676-697.

Maljkovic, V. & Nakayama, K. (1994). Priming of popout: I. Role of features. *Memory & Cognition*, 22(6), 657-672.

Maljkovic, V. & Nakayama, K. (2000). Priming of popout: III. A short-term implicit memory system beneficial for rapid target selection. *Visual Cognition*, 7 (5), 571–595.

Morison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming saccades. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 667-682

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Science* , 11, 520-527.

Potter, M. C. (1999). Understanding Sentences and Scenes: The Role of Conceptual Short-Therm Memory. In: V. Coltheart, *Fleeting Memories: Cognition of Brief Visual Stimuli* (pp. 13-43). Cambridge, Massachusetts: The MIT Press.

Potter, M. C. (1993). Very short-therm conceptual memory. *Memory & Cognition* , 2 (21), 156-161.

Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, 331, 163-166.

Rossini, J. C., & Galera, C.(2008). Seleção e Análise de Estímulos na Tarefa de Busca Visual. *Psicologia: Reflexão e Crítica*, 21(1), 20-27

Rousselet, G. A., Fabre-Thorpe, M., & Thorpe, S. J. (2002). Parallel processing in high-level categorization. *Nature Neuroscience* (5), 629-630.

Rumelhart, D. E. & Ortony, A. (1977). The representation of knowledge in memory. In: R. C. Anderson, R. J. Spiro & W. E. Montague (Ed.), *Schooling and the acquisition of knowledge*. (pp. 99-135). Hillsdale: Erlbaum.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology* , 16, 97-136.

Treisman, A. (1988). Features and objects: The fourteenth Barlett memory lecture. *Quarterly Journal of Experimental Psychology:human Experimental Psychology*, 40(A), 201-237.

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology* , 59 (11), 1931-1949.

Underwood, G., Templeman, E., Lamming, L., & Foulshan, T. (2008). Is attention necessary for object identification? Evidence from eye movements during the inspection of real-world scenes. *Consciousness and Cognition* , 17 (1), 159-170.

Wolfe, J. M. (2007). Guided Search 4.0: Current Progress with a model of visual search. In: W. G. (Ed.), *Integrated Models of Cognitive System* (pp. 99-119). New York: Oxford.

Wolfe, J. M. (1998). Visual Search. In: H. Pashler, *Attention*.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided Search: An alternative to feature integration model for visual search. *Journal of Experimental Psychology:Human Perception and Performance* , 15 (3), 419-433.

Wolfe, J. M., Oliva, A., Horowitz, T. S., Butcher, S. J. & Bompas, A. (2002). Segmentation of objects from backgrounds in visual search tasks. *Vision Research*, 42, 2958-3004.

Aceite do CEP

Realçar texto



Universidade Federal de Uberlândia

Pró-Reitoria de Pesquisa e Pós-Graduação

COMITÊ DE ÉTICA EM PESQUISA - CEP

Avenida João Naves de Ávila, nº. 2160 - Bloco J - Campus Santa Mônica - Uberlândia-MG -

CEP 38400-089 - FONE/FAX (34) 3239-4131

e-mail: cep@propp.ufu.br; www.comissoes.propp.ufu.br

ANÁLISE FINAL Nº. 245609 DO COMITÊ DE ÉTICA EM PESQUISA PARA O PROTOCOLO REGISTRO
CEP/UFU 085/09

Projeto Pesquisa: O processamento da informação semântica na busca visual: Uso de dicas
endógenas como textura na tarefa de busca visual.

Pesquisador Responsável: Joaquim Carlos Rossini

De acordo com as atribuições definidas na Resolução CNS 196/96, o CEP manifesta-se pela aprovação
do projeto de pesquisa proposto.

O protocolo não apresenta problemas de ética nas condutas de pesquisa com seres humanos, nos limites
da redação e da metodologia apresentadas.

O CEP/UFU lembra que:

a- segundo a Resolução 196/96, o pesquisador deverá arquivar por 5 anos o relatório da pesquisa e os
Termos de Consentimento Livre e Esclarecido, assinados pelo sujeito de pesquisa.

b- poderá, por escolha aleatória, visitar o pesquisador para conferência do relatório e documentação
pertinente ao projeto.

c- a aprovação do protocolo de pesquisa pelo CEP/UFU dá-se em decorrência do atendimento a
Resolução 196/96/CNS, não implicando na qualidade científica do mesmo.

Data de entrega do relatório final: setembro de 2010.

SITUAÇÃO: PROTOCOLO DE PESQUISA APROVADO

O CEP/UFU LEMBRA QUE QUALQUER MUDANÇA NO PROTOCOLO DEVE SER INFORMADA
IMEDIATAMENTE AO CEP PARA FINS DE ANÁLISE E APROVAÇÃO DA MESMA.

Uberlândia, 17 de junho de 2009.

Prof. Dra. Sandra Terezinha de Farias Furtado
Coordenadora do CEP/UFU

Orientações ao pesquisador

- O sujeito da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado (Res. CNS 196/96 - Item IV.1.f) e deve receber uma cópia do Termo de Consentimento Livre e Esclarecido, na íntegra, por ele assinado (Item IV.2.d).
- O pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado e descontinuar o estudo somente após análise das razões da descontinuidade pelo CEP que o aprovou (Res. CNS Item III.3.z), aguardando seu parecer, exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade de regime oferecido a um dos grupos da pesquisa (Item V.3) que requeiram ação imediata.
- O CEP deve ser informado de todos os efeitos adversos ou fatos relevantes que alterem o curso normal do estudo (Res. CNS Item V.4). É papel de o pesquisador assegurar medidas imediatas adequadas frente a evento adverso grave ocorrido (mesmo que tenha sido em outro centro) e enviar notificação ao CEP e à Agência Nacional de Vigilância Sanitária - ANVISA - junto com seu posicionamento.
- Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas. Em caso de projetos do Grupo I ou II apresentados anteriormente à ANVISA, o pesquisador ou patrocinador deve enviá-las também à mesma, junto com o parecer aprobatório do CEP, para serem juntadas ao protocolo inicial (Res.251/97, item III.2.e). O prazo para entrega de relatório é de 120 dias após o término da execução prevista no cronograma do projeto, conforme norma.