Universidade de São Paulo
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto
Departamento de Psicologia

Leonardo Gomes Bernardino

Tamanhos pictóricos percebidos sobre gradientes de textura desenhados e fotografados após exposições breves

Ribeirão Preto

2012

Leonardo Gomes Bernardino

Tamanhos pictóricos percebidos sobre gradientes de textura desenhados e fotografados após exposições breves

Tese apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo, como parte das exigências para obtenção do título de Doutor em Ciências.

Área de Concentração: Psicobiologia.

Orientador: Sérgio Sheiji Fukusima

Ribeirão Preto

2012

Autorizo a reprodução e/ou divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

FICHA CATALOGRÁFICA

Bernardino, Leonardo Gomes

Tamanhos pictóricos percebidos sobre gradientes de textura desenhados e fotografados após exposições breves. Ribeirão Preto, 2012.

143 p.: il.; 30 cm.

Tese de Doutorado apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto/USP - Departamento de Psicologia. Área de concentração: Psicobiologia.

Orientador: Fukusima, Sérgio Sheiji

1. Percepção de tamanho. 2. Gradiente de textura. 3. Tempo de exposição. 4. Mecanismos *bottom-up* e *top-down*. 5. Movimentos oculares

FOLHA DE APROVAÇÃO

Leonardo Gomes Bernardino

Tamanhos pictóricos percebidos sobre gradientes de textura desenhados e fotografados após exposições breves

Tese apresentada à Faculdade de Filosofia,
Ciências e Letras de Ribeirão Preto da
Universidade de São Paulo para obtenção do
título de Doutor em Ciências.
Área de Concentração: Psicobiologia.

Aprovado em: ____/____/____

Banca Examinadora

Prof. Dr. Sérgio Sheiji Fukusima (Orientador)

Instituição: FFCLRP - Universidade de São Paulo Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Para Mirian, porto, vela e buriti

AGRADECIMENTOS

Sou grato a muitos nesta minha trajetória. Do ponto de vista científico, agradeço ao meu orientador e professor Sérgio Sheiji Fukusima, por me ensinar a ser um pesquisador crítico e ético e por ter paciência com a minha teimosia. Ao professor Jose Antonio Aznar Casanova, que me acolheu em um país estrangeiro com simpatia, amizade e me ensinou que não há limites. Aos professores César Alexis Galera, Nelson Torro Alves e José Aparecido da Silva, pelos comentários e sugestões valiosas no exame de qualificação. Ao Igor, pelo auxílio técnico. A Renata Vicentini, Regina Teles e Maria Inês Joaquim, pela disposição em atender minhas solicitações na pós-graduação. As colegas do laboratório de Percepção e Psicofísica, Cris e Patrícia, por dividirem comigo suas vidas e carreiras científicas. Aos colegas do departamento de Psicologia Básica da Universidade de Barcelona, Hector, Cris, Manuel e Mathias, por todo o apoio. Aos amigos de pós-graduação, Laranja, Ricardo, Rafael, Dudu, Thiago, Roberto, Lucas, Milena, Mikael e Édson, pelos momentos de seriedade e de diversão. A todos os que me concederam seu tempo e participaram dos experimentos. A CAPES e ao CNPq, pelas bolsas concedidas.

No entanto, nestes quatro anos não fui apenas cientista ou pesquisador. Essa é apenas uma de minhas facetas, que afeta e é influenciada pelo o que sou, vejo, penso e vivo. Assim, não posso deixar de agradecer a Deus, princípio e fim e a Virgem Maria, que me preservou por onde andei. Aos meus pais, Vilmar e Irinéa, os responsáveis por tudo o que eu sou, que me criaram com carinho, amor e me mostraram o caminho, mais com exemplos do que com palavras. A Larissa, irmã e amiga, mulher forte e decidida, sempre uma referência e quem eu tanto admiro. Ao João Vítor, que agora poderá ler este agradecimento, pelas melhores conversas que tive recentemente. Ao Bruno, amigo e irmão, com quem compartilhei a casa e o laboratório, com quem a sintonia é tão boa, que eu não sei se uma ideia foi minha ou dele. Aos grandes parceiros, Hugo, Rui e Fantim, as terças nunca mais serão as mesmas. Aos amigos Adriele, Edu, Luiza, Steve, Pol e Gabrielle, minha família em Barcelona. Aos primos, Ita, Raoni e Vinícius, por esticarem o meu tempo. As amigas, Sol, Thê, Carol e Lina, por sempre estarem onde eu preciso. A todos os meus familiares, por torcerem pelo meu sucesso. A minha grande família holandesa, por me acolherem com tanto zelo.

A todos, muito obrigado!

“A gente quer passar um rio a nado, e passa;
mas vai dar na outra banda é num ponto
muito mais embaixo, bem diverso do em que
primeiro se pensou. Viver nem não é muito
perigoso?”

João Guimarães Rosa

RESUMO

Bernardino, L. G. (2012). **Tamanhos pictóricos percebidos sobre gradientes de textura desenhados e fotografados após exposições breves**. 143 p. Tese (Doutorado) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto-SP, 2012.

O presente estudo teve por objetivo investigar a dinâmica temporal da percepção de tamanho de objetos inseridos em gradientes de textura, provenientes de desenho gerado computacionalmente (gradiente de perspectiva) e de fotografia com elementos naturais (gradiente fotografado). Além disso, verificou-se a ocorrência e padrão dos movimentos oculares nestas condições. Para isso, foram realizados 2 experimentos. No primeiro experimento, dois círculos negros eram apresentados no meridiano vertical e a tarefa dos 96 participantes foi indicar em qual parte do campo visual, superior ou inferior, foi apresentado o maior círculo. Usando o método das escadas duplas, estes círculos eram apresentados brevemente (50, 100, 150 ou 200 ms) em três condições de fundos de tela (sem textura, gradiente de perspectiva ou gradiente fotografado). Foram calculados a inclinação da função psicométrica e o ponto de igualdade subjetivo (PIS) como medidas da sensibilidade discriminativa e da distorção perceptiva, respectivamente. Os resultados mostraram que os participantes discriminaram melhor o tamanho dos estímulos à medida que o tempo de exposição aumentava, independentemente do gradiente de textura. A análise do PIS revelou que há uma forte distorção do tamanho aos 100 ms para o gradiente de perspectiva e aos 150 ms para o gradiente fotografado. No segundo experimento, 24 participantes realizaram uma tarefa semelhante ao do primeiro experimento e utilizou-se um *eye tracker* para registrar o movimento ocular. Os resultados mostraram que, em apresentações de 150 e 200 ms, os movimentos oculares ocorreram em menos de 10% das tentativas em todos os fundos de tela. Isto indica que a redução das distorções de tamanho após os 100 ms observada no primeiro experimento não é explicada completamente pela alocação da atenção. Tomados em conjunto, estes resultados sugerem que o processamento e integração das informações de tamanho e profundidade são mediados principalmente por mecanismos *bottom-up*. Ademais, em uma condição de menor restrição temporal (1500 ms), a presença de informações de profundidade afetou a dimensão espacial das fixações e sacadas, mas não a dimensão temporal. Os participantes olharam preferencialmente para o estímulo apresentado no campo visual superior, o que indica que a atenção foi alocada às posições de maior profundidade. Os resultados também indicaram que os estímulos maiores capturaram tanto a atenção transitória quanto a atenção sustentada. Assim, este estudo permitiu um maior conhecimento sobre os mecanismos perceptivos e atentos envolvidos no processamento das informações de tamanho e de profundidade.

Palavras-chave: Percepção de tamanho. Gradiente de textura. Tempo de exposição. Mecanismos *bottom-up* e *top-down*. Movimentos oculares.

ABSTRACT

Bernardino, L. G. (2012). **Pictorial size perceived under line-drawing and photographed texture gradient after brief exposures**. 143 p. Tese (Doutorado) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto-SP, 2012.

This study aimed to investigate the temporal dynamics of object size perception under texture gradients from line-drawing (perspective gradient) and photograph with natural elements (photographed gradient). Additionally, it was analyzed the occurrence and the pattern of eye movements in these conditions. For this, we carried out 2 experiments. In the first experiment, two black circles were displayed in the vertical meridian and 96 participants reported whether the bigger one was presented at the lower or upper visual field. By using double staircase psychophysics method, these circles were briefly presented (50, 100, 150 or 200 ms) in three background conditions (no texture, perspective gradient or photographed gradient). The slope of the psychometric function and the points of subjectivity equality (PSE) were calculated and used as discrimination sensitivity and size distortion measures, respectively. The results showed a greater sensitivity to size discrimination as exposure time increased, regardless the texture gradient. The analysis of PSE indicated greater size distortions in perspective gradient at 100 ms and in photographed gradient at 150 ms. In the second experiment, 24 participants performed a task similar to that in Experiment 1 while an eye tracker recorded their eye-movements. The results showed that, at 150 and 200 ms, eye movements occurred in less than 10% of trials in all backgrounds. It indicates that the size distortion reduction after 100 ms observed in the first experiment cannot be fully explained by attention allocation. Taken together, these results suggest that the process and integration of size and depth information are mainly mediated by *bottom-up* mechanism. Moreover, in a less time constrain condition (1500 ms), depth cues affected the spatial measures of fixations and saccades, but not the temporal ones. The participants looked more frequently at the stimulus on the upper visual field, indicating that attention has been allocated to positions of greater depth. The results also indicated that large stimulus capture both transient and sustained attention. Thus, this study provides a better understanding of perceptual and attentive mechanisms involved in size and depth process.

Keywords: Size perception. Texture gradient. Exposure time. Bottom-up and top-down mechanisms. Eye movements.

LISTA DE ILUSTRAÇÕES

Figura 1 - Exemplos de estímulos na condição controle (a), na condição com gradiente de perspectiva (b) e na condição com gradiente fotografado (c). Em todos os exemplos, ambos os círculos têm o mesmo tamanho (100 pixels no tamanho original).....	44
Figura 2 - Inclinação média das funções psicométricas para comparações de tamanho realizadas nas 3 condições de fundo de tela (controle, gradiente de perspectiva e gradiente fotografado) e nos 4 tempos de exposição (50, 100, 150 e 200 ms).....	48
Figura 3 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado) e nos 4 tempos de exposição (50, 100, 150 e 200 ms).....	49
Figura 4 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas na condição controle nos 4 tempos de exposição (50, 100, 150 e 200 ms)	50
Figura 5 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas no gradiente de perspectiva nos 4 tempos de exposição (50, 100, 150 e 200 ms).....	51
Figura 6 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas no gradiente fotografado nos 4 tempos de exposição (50, 100, 150 e 200 ms).....	52
Figura 7 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas aos 50 ms nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado)	53
Figura 8 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas aos 100 ms nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado)	54
Figura 9 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas aos 150 ms nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado)	55
Figura 10 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas aos 200 ms nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado)	56
Figura 11 - Foto do equipamento e do sistema de duas telas (dual screen) utilizado para a execução do experimento nas condições breves de apresentação dos estímulos (150 e 200 ms). O <i>eye tracker</i> SMI iView x RED500 está localizado abaixo do monitor de 22 polegadas	64

- Figura 12** - Um esquema das áreas de interesse (AOIs) estabelecidas para a análise dos parâmetros oculares. A área 1 é o ponto de fixação (PF), a área 2 é o círculo apresentado acima (CAc), a área 3 é o círculo apresentado abaixo (CAb) e a área 4 é o fundo de tela (F). 67
- Figura 13** - Porcentagem dos erros de julgamentos nas comparações de tamanho nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado) e nos 3 tempos de exposição (150, 200 e 1500 ms). 69
- Figura 14** - Duração total média das fixações e dos movimentos sacádicos (em ms) em cada uma das 4 áreas de interesse (PF, CAc, CAb e F), em função dos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado)..... 78

LISTA DE TABELAS

Tabela 1 - O número de erros de julgamentos dos participantes nos 3 tempos de exposição (150, 200 e 1500 ms), em função dos 11 tamanhos do estímulo teste.	70
Tabela 2 - A porcentagem (%) de tentativas em que ocorreram movimentos oculares, em função dos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado) e dos 2 tempos de exposição (150 e 200 ms).....	71
Tabela 3 - O número total de fixações dos participantes para os 11 tamanhos do estímulo teste, em função dos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado).	73
Tabela 4 - O número total de fixações dos participantes nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado), em função das 4 áreas de interesse (PF, CAC, CAb e F).	75
Tabela 5 - O número total de fixações dos participantes nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado), em função do tamanho dos círculos (maior e menor).	75
Tabela 6 - A frequência da localização espacial das fixações nas 4 áreas de interesse (PF, CAC, CAb e F) para os 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado), em função da sequência em que ocorreram até a sexta fixação.	77
Tabela 7 - O número total de vezes em que o primeiro movimento sacádico foi dirigido acima ou abaixo no campo visual, em função dos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado).....	80
Tabela 8 - O número total de vezes em que o primeiro movimento sacádico foi dirigido ao maior ou ao menor círculo, em função dos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado).....	81

LISTA DE ABREVIATURAS E SIGLAS

ANOVA: análise de variância

AOI: área de interesse

Cac: área de interesse correspondente ao círculo acima

Cab: área de interesse correspondente ao círculo abaixo

cd/m²: candela por metro quadrado

cm: centímetro

EPM: erro padrão da média

HITD: hipótese da invariância tamanho-distância

Hz: Hertz

PF: área de interesse correspondente ao ponto de fixação

F: área de interesse correspondente ao fundo de tela ou gradiente de textura

mm: milímetro

ms: milissegundo

n: número total

NGL: núcleo geniculado lateral

PIS: ponto de igualdade subjetiva

V1: área visual primária

$\bar{x}$: média amostral

χ^2 : distribuição de qui-quadrado

SUMÁRIO

1. Introdução	16
1.1 Preâmbulo	17
1.2 Percepção de tamanho e percepção de distância	18
1.3 Gradiente de textura	22
1.4 Dinâmica temporal da percepção de tamanho e profundidade	26
1.5 Atenção: fator modulador da percepção de tamanho e profundidade	30
1.6 Movimentos oculares: mecanismos <i>bottom-up</i> e <i>top-down</i>	34
2. Objetivos gerais	40
3. Experimento 1	42
3.1 Objetivo	43
3.2 Método	43
3.2.1 Participantes	43
3.2.2 Materiais e equipamentos	43
3.2.3 Procedimento	45
3.2.4 Análise dos dados	46
3.3 Resultados	47
3.3.1 Inclinação da função psicométrica	47
3.3.2 Ponto de igualdade subjetivo (PIS)	48
3.4 Discussão	56
4. Experimento 2	62
4.1 Objetivo	63
4.2 Método	63
4.2.1 Participantes	63
4.2.2 Materiais e equipamentos	63
4.2.3 Procedimento	65
4.2.4 Análise dos dados - Respostas	66
4.2.5 Análise dos dados - Parâmetros oculares	66
4.3 Resultados	68
4.3.1 Erros de julgamento	68
4.3.2 Tempos de exposição de 150 e 200 ms	70
4.3.3 Tempo de exposição de 1500	72
4.3.3.1 Tempo de resposta	72
4.3.3.2 Número de fixações	72
4.3.3.3 Dimensão espacial das fixações	74
4.3.3.4 Sequência das fixações	76
4.3.3.5 Dimensão temporal das fixações	77
4.3.3.6 Dimensão espacial dos movimentos sacádicos	79
4.3.3.7 Amplitude dos movimentos sacádicos	81
4.4 Discussão	82
5. Discussão geral	87
6. Conclusão	94

Referências	96
Apêndices	112
APÊNDICE A – Artigo apresentado no exame de qualificação	113
Anexos	135
ANEXO A – Termo de Consentimento Livre e Esclarecido	136
ANEXO B – Parecer do Comitê de Ética em Pesquisa da FFCLRP - USP	137
ANEXO C – Instruções aos participantes – Experimento 1	138
ANEXO D – Normas do Comitê de Ética da Universidade de Barcelona	139
ANEXO E – Instruções aos participantes – Experimento 2	142

1. Introdução

1.1. Preâmbulo

Imagine que um dia, sem razão aparente, você pare de enxergar. O que resta é um clarão branco à sua frente. Paulatina e sucessivamente, todas as pessoas ao seu redor também perdem a capacidade de ver, embora os outros sentidos permaneçam intactos. Quais seriam as repercussões desta insidiosa incapacidade repentina? Esta não é uma ideia original. O escritor português José Saramago, ganhador do prêmio Nobel de Literatura em 1998, desenvolve este tema em seu livro *Ensaio sobre a Cegueira* (1995), o qual foi adaptado ao cinema por Fernando Meirelles em 2008. Não é preciso conhecer o livro ou o filme, basta imaginar e a conclusão será semelhante: as consequências seriam catastróficas.

Por volta dos 9 anos de idade, todos os dias eu andava de bicicleta pelo bairro onde eu morava. Um dia, fui desafiado por um amigo: duvido que você passe por esta rua de olhos fechados. Eu era criança e não podia escutar a palavra “duvido”, esta era a senha para realizar qualquer coisa. Além disso, eu conhecia muito bem aquela rua e julguei não haver riscos. O resultado, é claro, foi catastrófico também: bati em um senhor que andava de bicicleta, estragando ambas as bicicletas e quase quebrei o nariz.

Este pequeno preâmbulo teve por objetivo sublinhar que somos uma espécie eminentemente visual. Por isso, a visão foi o sentido que mais atraiu a atenção dos pesquisadores ao longo do tempo e esta prevalência ainda continua. Tal observação não implica que os outros sentidos (audição, gustação, olfação e somestesia) desempenham um papel marginal em nossa interação com o ambiente ou que sejam menos complexos. Inclusive, dada a complexidade dos sistemas gustativo e olfativo, cujos receptores são químicos, nossos conhecimentos sobre estes ainda são poucos.

De maneira geral, os sistemas sensoriais apresentam uma organização mista, hierárquica e em paralelo para o processamento dos estímulos nas etapas conhecidas como transdução e codificação. Na primeira, é realizada a absorção da energia do estímulo ambiental, a qual é transformada em um potencial biológico gerado pelas membranas dos receptores. Em seguida, na fase de codificação, este potencial é transformado em potenciais de ação, que permitem sua transmissão e a continuidade do processamento em outros níveis do sistema nervoso central. Cabe ressaltar que cada sistema sensorial apresenta uma

modulação própria para o reconhecimento consciente do estímulo (Kandel, Schwartz & Jessel, 2000).

No caso da visão, o olho humano e suas estruturas captam a luz presente no ambiente e a direcionam ao fundo do olho, onde está localizada a retina. Esta estrutura contém os fotorreceptores (cones e bastonetes) que realizam a transdução sensorial e transmitem estes sinais elétricos ao cérebro através do nervo óptico. À medida que estes sinais alcançam níveis hierárquicos mais complexos no sistema nervoso central, as informações que eles carregam são ordenadas, processadas e interpretadas, gerando de maneira contínua nossa percepção visual do mundo que nos circunda.

Assim, o tempo todo nosso sistema visual realiza estimativas de velocidade, tamanho e distância, o que nos permite interagir com objetos e pessoas ao nosso redor, tomar decisões, planejar e executar ações. Dessa maneira, alcançar uma xícara de café sobre uma mesa, jogar um papel no cesto de lixo ou dirigir um carro são exemplos de tarefas cotidianas mediadas pelo sistema visual e que demandam a realização e atualização constantes destas estimativas, mesmo que não estejamos conscientes de realizá-las. Pode-se afirmar, portanto, que a percepção do espaço está orientada para a ação e que, neste contexto, as informações de tamanho e de distância (profundidade) desempenham um papel fundamental. Assim, o presente trabalho teve por objetivo verificar o tempo necessário para processar e integrar as informações de tamanho e de profundidade, investigando aspectos perceptivos e atentos implicados neste processo.

1.2. Percepção de tamanho e percepção de distância

O tamanho percebido de um objeto depende do ângulo visual que este subtende na retina. Ou seja, se um objeto subtende um ângulo maior do que outro objeto, logo será percebido como maior que este. Além disso, do ponto de vista neural, há evidências de que no córtex temporal há neurônios cuja resposta é modulada em função do tamanho dos objetos e outros que respondem seletivamente a determinado tamanho (Rousset, Thorpe e Fabre-Thorpe, 2004a). A partir destes argumentos, pode-se concluir erroneamente que este é um tema cujo estudo tenha se esgotado e que não lance novos desafios aos pesquisadores. Esta é

uma conclusão deveras equivocada, pois o tamanho dos objetos depende e é modulado por uma série de fatores.

Inicialmente, pode-se destacar que o ângulo visual subtendido de um objeto na retina depende de dois fatores: o tamanho real do objeto e a distância deste para o observador. Isto é, quanto mais distante um objeto, menor será o tamanho projetado deste na retina. Assim, as informações de tamanho e de distância são inversamente proporcionais e estão intimamente relacionadas. Contudo, essa relação não é tão direta e inequívoca como pode parecer. O tamanho percebido altera-se não linearmente (muda menos) em relação ao tamanho da imagem projetado na retina à medida que aumenta a distância do observador (Foley, Ribeiro-Filho & Da Silva, 2004).

Também há evidências de que não temos acesso direto à informação retiniana do tamanho. Por exemplo, em um estudo psicofísico encontrou-se que a representação do tamanho não está exclusivamente condicionada ao tamanho retinal, dependendo também do tamanho percebido (Found & Muller, 2001). Desta maneira, pode-se concluir que a informação de distância é utilizada na percepção de tamanho. Em outro estudo, Sperandio, Savazzi, Gregory e Marzi (2009) mostraram que o tempo de reação diminui em uma maior proporção com o aumento do tamanho percebido do que com o aumento do tamanho retiniano. Estes achados são corroborados por estudos neurofisiológicos, os quais indicam que a atividade do córtex visual primário (V1) reflete não o tamanho do *input* retiniano, mas sim o tamanho percebido (Murray, Boyaci & Kersten, 2006) e que V1 encontra-se ativo durante a observação de uma ilusão visual de tamanho (Macevoy & Fitzpatrick, 2006).

Ao longo do século passado, inúmeros pesquisadores se lançaram ao desafio de explicar a relação entre a percepção de tamanho e a percepção de distância. Em um artigo de revisão publicado há mais de 50 anos (Epstein, Park & Casey, 1961) foram listadas as seguintes hipóteses para explicar tal relação: 1) Hipótese da Invariância Tamanho-Distância (HITD); 2) Hipótese do Tamanho Conhecido-Distância Aparente¹; e 3) Lei de Emmert². Dentre estas, a mais amplamente aceita é a HITD, a qual postula que o ângulo visual é transformado em tamanho percebido após a distância ser registrada (Kilpatrick & Ittelson, 1953).

Esta hipótese foi comprovada no clássico estudo realizado por Holway e Boring (1941) nos corredores do departamento de Psicologia da Universidade de Harvard. Um fato curioso deste experimento é que ele foi realizado durante a madrugada para que a incidência

¹ Esta hipótese prevê que com a informação de que o tamanho do objeto é constante, mudanças discretas no tamanho da imagem retinal, levará à percepção de mudança na distância aparente.

² Esta hipótese surgiu de investigações de pós-imagem, as quais teriam seu tamanho de projeção diretamente proporcional à distância dos olhos para a superfície de projeção. É considerada um caso especial da HITD.

de luz fosse a menor possível. Neste estudo, a tarefa dos observadores foi ajustar o diâmetro de um círculo de comparação ao diâmetro real do círculo padrão. O experimento consistiu em 4 fases diferentes, em função da quantidade de informações de profundidade disponíveis, a saber: 1) condição visual binocular, 2) condição visual monocular, 3) condição visual monocular através de uma pequeníssima abertura (*pinhole*³) e 4) condição visual monocular através de um *pinhole* acoplado a extremidade de um túnel de redução (as paredes do corredor foram cobertas com panos negros). Os resultados revelaram que à medida que as informações sobre distância foram progressivamente reduzidas, o tamanho percebido aproximava-se bastante do ângulo visual do objeto.

Embora seja a mais aceita, a HITD não explica todos os fenômenos implicados na relação tamanho-distância. O exemplo mais famoso talvez seja o da ilusão da Lua, na qual a Lua próxima ao horizonte parece bem maior do que quando localizada no zênite. De acordo com as predições da HITD seria esperado exatamente o oposto, ou seja, objetos mais distantes (neste caso, o zênite) seriam percebidos como maiores. Este fenômeno é conhecido como o paradoxo tamanho-distância (Gruber, 1954). Além disso, o paradoxo tamanho-distância é considerado como uma evidência de que a percepção de tamanho afeta a percepção de distância, o que não é corroborado pelo estudo de Mon-Williams e Tresilian (1999). Estes autores afirmam que tal paradoxo é explicado por componentes cognitivos envolvidos neste processo, o que reduziria a influência da informação sensorial propriamente dita.

Na tentativa de explicar porque certos fenômenos relativos à relação tamanho-distância escapam às predições da HITD, há autores que defendem que o processamento das informações de tamanho e distância ocorre em diversos níveis (por exemplo, Higashiyama, 1979), preservando assim alguns pontos da hipótese em questão. Neste contexto, uma elegante teoria foi proposta por Gogel e Da Silva (1987), separando os fatores perceptuais e os fatores cognitivos envolvidos no processamento de tamanho e distância. Haveria, portanto, dois níveis de processamento para se estimar tamanho e distância: 1) perceptivo, que seria acessado através de indícios imediatamente disponíveis ao observador; e 2) cognitivo, que seria dependente do primeiro, porém mediado pela memória ou experiências passadas. Em um estudo mais recente, sugere-se que os fatores cognitivos são tão relevantes, que não há diferença entre se estimar o tamanho de um objeto valendo-se da visão ou da memória, em termos de tamanho familiar (Haber & Levin, 2001).

³ O *pinhole* aumenta o tamanho da imagem retiniana de objetos próximos, o que minimiza os efeitos da acomodação, que é uma informação de profundidade oculomotora.

Por outro lado, há autores que defendem a independência total no processamento das informações de tamanho e distância (ver Day & Parks, 1989), ideia esta bastante controversa e difícil de ser sustentada. No entanto, nos últimos anos aumentaram as evidências de que o processamento da informação de tamanho requer conhecimento sobre a distância percebida do objeto, porém o contrário não é verdadeiro (Tozawa & Oyama, 2006). Desta maneira, a percepção de tamanho e a percepção de distância seriam processos independentes em algum nível, mas não totalmente. O sistema visual, portanto, dissociaria o processamento de tamanho e distância em função da quantidade de informações disponíveis e confiáveis em um determinado momento. Neste sentido, O'Brien (1997) argumenta que tamanho e distância são independentemente processados em condições sem nenhum tipo de restrição visual, porém esta observação não é verdadeira reduzindo-se a quantidade de indícios de distância, por exemplo, em condição visual monocular ou em tempos breves de exposição. Neste caso, utiliza-se a relação entre tamanho e distância para forjar uma percepção mais acurada do espaço.

Outra teoria clássica para explicar a relação entre tamanho e distância é a da abordagem direta proposta por Gibson (1950). Este modelo também é conhecido como abordagem ecológica, pois salienta o caráter natural e adaptativo da percepção de uma maneira espécie-dependente. Este autor defendia que as imagens que são formadas em nossa retina possuem toda a informação necessária para o registro do ambiente e das características dos objetos e das relações entre eles. Assim, uma mediação cognitiva e o envolvimento de altos níveis na hierarquia da circuitaria cerebral (*top-down*) não seriam fundamentais para perceber o mundo ao nosso redor.

Cabe ressaltar a importância dada por Gibson às alterações na textura das superfícies (gradiente de textura), consideradas um fator crítico para perceber o espaço em geral e a distância e o tamanho especificamente. Seguindo esta ideia, o cômputo da primeira é irrelevante para a representação do segundo. Recentemente, Bian e Andersen (2010) investigaram o processamento de objetos dispostos no solo (*ground*) e no teto (*ceiling*) em posição natural e invertida. Os resultados revelaram o papel fundamental da superfície na organização de nossa representação do mundo. Tais achados são consonantes com o modelo gibsoniano, embora os autores indiquem o envolvimento de mecanismos *top-down* nesta organização. Neste sentido, ainda há evidências de que a percepção de *ground* prevalece sobre a de *ceiling* numa condição de rivalidade binocular (Ozkan & Braunstein, 2009).

Desde o clássico estudo de Gibson (1950), o gradiente de textura é indicado como uma condição necessária e suficiente para o cálculo de distância entre o observador e um objeto (distância egocêntrica) ou entre dois objetos (distância exocêntrica). Assim, na próxima seção serão apresentadas investigações sobre o gradiente de textura, enfocando sua influência sobre a percepção de tamanho.

1.3. Gradiente de textura

Compreender como nosso sistema visual representa o ambiente em três dimensões a partir de imagens bi-dimensionais formadas na retina é uma aspiração antiga dos estudiosos da percepção visual do espaço. A resolução desse problema é possível graças aos indícios pictóricos de profundidade, que são informações contidas no ambiente e que podem ser apreendidas em condição monocular (um olho somente) ou binocular (ambos os olhos). Dentre os indícios monoculares, pode-se citar a oclusão, o tamanho relativo, a altura no campo visual, a perspectiva aérea, a paralaxe do movimento, o sombreamento e o gradiente de textura. Por outro lado, a convergência, a acomodação e a disparidade binocular são indícios binoculares (Cutting & Vishton, 1995).

Desse modo, o sistema visual constrói uma representação geral de superfícies em espaço tridimensional, primeiramente a partir de indícios de profundidade individuais monoculares e binoculares, que são integrados posteriormente. Nos últimos anos, muitos estudos foram realizados com o objetivo de explicar como ocorre essa integração dos indícios de profundidade (Knill & Saunders, 2003; Oruç, Maloney & Landy, 2003; Poom & Borjesson, 1999).

Apesar do crescente número de publicações e dos esforços nesse sentido, o modo como os indícios afetam individualmente a percepção de profundidade ainda não está completamente elucidado. Confirmando as afirmações de Gibson (1950), Sinai, Ooi e He (1998) mostraram que, quando há informação disponível sobre a textura da superfície, a percepção de distância absoluta é bastante acurada. Assim, qualquer interrupção ou incongruência no gradiente diminui significativamente a precisão nos julgamentos de distância realizados pelos observadores.

Uma boa e simples descrição sobre o gradiente de textura é apresentada por Schiffman (2005): é uma informação pictórica monocular de profundidade, na qual à medida que aumenta a distância para o observador, os elementos constituintes da superfície parecem menores e projetam texturas mais finas ou densas. Entretanto, esse conceito é assaz amplo e não apresenta uma importante característica deste indício de profundidade.

A informação contida no gradiente de textura pode ser decomposta em três: *perspective* (escalonamento), *foreshortening* (compressão) e *density* (densidade). O escalonamento é a mudança no tamanho projetado dos elementos de textura, a compressão consiste em alterações na forma projetada destes elementos e a densidade é a mudança no número de elementos por unidade de área (Hillis et al., 2004). O escalonamento também é chamado de gradiente de perspectiva (Cutting & Millard, 1984) ou de gradiente de tamanho (Cutting & Vishton, 1995).

Essas informações distintas presentes no gradiente de textura foram investigadas por Cutting e Millard (1984). Eles estudaram três gradientes (perspectiva, compressão e densidade) para a percepção de superfícies planas e curvas. Os resultados indicaram que os gradientes de perspectiva e de densidade foram os mais importantes para o julgamento de distância na superfície plana.

Muitos estudos posteriores corroboraram esses resultados, mostrando principalmente a grande efetividade do gradiente de perspectiva para informar distância, mesmo na presença de outras informações de textura ou de indícios binoculares (Allison & Howard, 2000; Goodenough & Gilliam, 1997; Loomis, Philbeck & Zahorik, 2002; Saunders & Backus, 2006; Wu, He & Ooi, 2007). Esses dados evidenciam o quão eficiente é o sistema visual para extrair as informações de profundidade contidas na perspectiva linear.

Do ponto de vista ontogenético, há muitas evidências de que a capacidade de utilizar a perspectiva como indício de profundidade desenvolve-se nos primeiros meses de vida, provavelmente entre 6 e 7 meses de idade (Watson & Enns, 2012). Por exemplo, em uma tarefa de *preferential-reaching*, observou-se uma grande sensibilidade à perspectiva linear em crianças de 7 meses de idade, porém não nas de 5 meses (Hemker & Kavšek, 2010). Também há estudos que situam o início da utilização deste indício ainda mais cedo. Em 4 experimentos realizados com o auxílio de um rastreador de movimentos oculares (*eye tracker*), mostrou-se que crianças aos 5 meses de idade, mas não as de 4 meses, eram hábeis em utilizar as informações de perspectiva (Frichtel & Lécuyer, 2007).

Na década de 90, duas importantes investigações perscrutaram a influência dos gradientes de textura sobre a percepção de tamanho. No primeiro estudo, Williams e Enns (1996) investigaram se a adição de um contexto circundante afetava a ilusão horizontal-vertical, na qual uma linha vertical é percebida como maior do que uma linha horizontal a despeito de possuírem o mesmo tamanho físico. Os resultados sugerem que as distorções de tamanho observadas foram moduladas diferencialmente em função dos gradientes de textura em que os estímulos foram inseridos.

No segundo estudo, a tarefa dos participantes era localizar um alvo (cilindro) de tamanho único em meio a outros cilindros distratores, que eram ou maiores ou menores que o alvo. Estes estímulos estavam inseridos em um fundo de gradiente combinado com as informações de compressão e perspectiva. Observou-se que o tempo de reação foi maior, isto é, localizar o estímulo-alvo era mais difícil nas seguintes condições: 1) estímulo-alvo maior que os distratores e percebido como mais próximo ao observador, e 2) estímulo-alvo menor que os distratores e percebido como mais distante do observador (Aks & Enns, 1997). Pode-se notar que o grau de dificuldade da tarefa aumentou nas condições incongruentes com a HITD, ou seja, nas situações que simulavam o paradoxo tamanho-distância.

Pode-se especular que o padrão de resultados reportados por estes autores seria o inverso, caso houvesse a presença da linha do horizonte e os estímulos estivessem localizados acima deste. Tal hipótese baseia-se nos achados de Ozkan e Braunstein (2010). Numa série de 4 experimentos para se estudar os efeitos da linha do horizonte na percepção de tamanho relativo, estes autores encontraram que a relação entre tamanho e distância segue proximamente a HITD se os estímulos estão localizados abaixo ou sobre a linha do horizonte. No entanto, na condição em que os estímulos estavam acima da linha do horizonte, os resultados se ajustam melhor ao observado no paradoxo tamanho-distância.

Uma década depois dos trabalhos de Williams e Enns (1996) e Aks e Enns (1997), Torro-Alves e Fukusima (2007) pediram aos participantes para comparar o tamanho de dois estímulos inseridos em gradientes de textura. Estes eram apresentados até que o participante emitisse seu julgamento (tempo livre). Os resultados indicaram uma superestimação dos estímulos apresentados no campo visual superior, percebidos como mais distantes. A distorção de tamanho chegava até a 10% na situação com gradiente de perspectiva, confirmando a grande efetividade dessa informação de profundidade para afetar o julgamento de tamanho.

Ademais, o estudo de Farran, Whitaker e Patel (2009) também corrobora tais achados. Estes autores encontraram que a adição gradativa de indícios pictóricos monoculares de profundidade (altura no campo visual, perspectiva linear e textura, nesta ordem) afeta fortemente a discriminação de tamanho. Este processo é cumulativo, porém atinge um platô após a adição de duas dicas de profundidade. Os resultados revelam que a informação de profundidade é difícil de ser ignorada, requer pouca atenção focada e, provavelmente, parte dela é processada nos estágios iniciais da percepção visual.

É importante observar que os estudos supracitados (Aks & Enns, 1996; Farran et al., 2009; Torro-Alves & Fukushima, 2007; Williams & Enns, 1996), que investigaram a integração das informações de tamanho e os gradientes de textura foram realizados com estímulos pictóricos gerados computacionalmente (desenhos). Uma interessante possibilidade de investigação desse assunto é o uso de fotografias de cenas reais ou fotografias contendo elementos naturais para representar as dicas de profundidade, como o gradiente de textura. Em décadas passadas, alguns estudos utilizaram fotografias de cenas naturais para representar informações de distância (Gibson & Bridgeman, 1987; Smith, 1958; Todd & Akerstrom, 1987). De maneira geral, os resultados mostram que os participantes são capazes de extrair informações complexas de profundidade a partir de fotografias de cenas naturais, embora a profundidade percebida seja comprimida, de maneira sutil no plano horizontal e mais intensamente no plano sagital (Kraft, Patterson & Mitchell, 1986).

Segundo Watanabe (2006), representar o espaço através de fotografias é equivalente a representá-lo por meio de desenhos gerados no computador, pois ambas encerram propriedades não-euclidianas (hiperbólicas) do espaço visualmente percebido. Além disso, outros estudos sugerem que a percepção de tamanho e a percepção de distância são afetadas da mesma maneira em fotografias e no ambiente real (Bengston, Stergios, Ward & Jester, 1980) e que as distorções do espaço observadas em ambos são similares (Hecht, van Doorn & Koenderink, 1999). Baseado nestas evidências, e dada sua maior validade ecológica, estudos que utilizem fotografias com informações de distância são fundamentais para estender a nossa compreensão acerca da representação do espaço pelo sistema visual.

A dinâmica temporal inicial dos processos perceptivos em geral e da integração tamanho-distância em particular é assunto ainda a ser explorado. Desta maneira, estudos adicionais empregando fotografias com informações de profundidade são importantes, principalmente com estímulos apresentados brevemente. Nossa hipótese é a de que em tempos

curtos, um gradiente de perspectiva gerado no computador será mais eficientemente processado do que um gradiente fotografado com elementos naturais devido à maior complexidade de informações contidas neste último.

1.4. Dinâmica temporal da percepção de tamanho e profundidade

Ao discorrer sobre a capacidade do sistema visual humano em representar o ambiente em três dimensões não se pode negligenciar a rapidez com que as informações são processadas. O sistema visual é capaz de extrair informações de uma cena complexa em 150 ms, como foi mostrado por Thorpe, Fize e Marlot (1996) e corroborado em estudos ulteriores (por exemplo, Rousselet, Thorpe & Fabre-Thorpe, 2004b). Neste breve tempo é possível, inclusive, determinar se um objeto estava presente ou não e até categorizá-lo em um nível básico.

Não há dúvidas de que a percepção visual de uma imagem, por exemplo, torna-se mais e mais sofisticada à medida que aumenta o tempo de observação desta. Décadas atrás, por exemplo, Gulick and Stake (1957) mostraram que a percepção de tamanho é severamente afetada por tempos breves de exposição. A partir da constatação que a visão é um processo inferencial, um maior tempo de inspeção permite a elaboração de hipóteses sobre o que é visto, que aumentam em complexidade com o acúmulo de informações ao longo do tempo (Reynolds, 1981). Subjacente a esta afirmação, encontra-se a ideia de que a visão é um processo hierárquico numa maneira *coarse-to-fine*, isto é, inicialmente há um processamento mais geral (global) seguido do processamento dos detalhes (local).

Os resultados descritos por Navon (1977) podem ser interpretados como evidência deste processamento *coarse-to-fine*. Ele apresentava uma letra grande (por exemplo, H) formada por pequenas letras, que poderiam ser iguais ou diferentes à letra grande (s, por exemplo). Os participantes reconheciam imediatamente a letra grande, sem que as letras pequenas afetassem esta identificação, independente se estas eram congruentes ou não com a letra grande. Por outro lado, a letra grande facilitava a identificação da menor caso fossem iguais. Conclui-se, portanto, que há uma forte prevalência no processamento da informação global em detrimento da informação local.

Do ponto de vista neural, também há evidências de que o sistema visual esteja organizado de uma maneira hierárquica, condizente com o modelo *coarse-to-fine*. Para melhor compreender esta afirmação, será apresentado um panorama sobre a transmissão da informação visual por meio de impulsos elétricos desde a retina até o córtex visual primário e deste para outras áreas superiores de processamento do sistema nervoso. Estas informações podem ser encontradas em textos básicos sobre neurociências em geral (por exemplo, Kandel et al., 2000) e sensação e percepção em específico (por exemplo, Goldstein, 2009).

Na retina humana, estrutura localizada no fundo do olho, encontra-se os fotorreceptores, células especializadas que codificam a informação luminosa em impulsos elétricos (transdução sensorial). Há dois tipos de fotorreceptores: os bastonetes e os cones. Os bastonetes são células alongadas e de formato cilíndrico, bastante numerosas (aproximadamente 100 milhões) e localizadas exclusivamente na periferia da retina, que são responsivas à baixa luminosidade, estando relacionadas à visão escotópica e à percepção de movimento. Os cones são células menores e mais estreitas, menos numerosas (cerca de 5 milhões) e altamente concentradas na fóvea (região central da retina), que são sensíveis à luz, permitindo a visão de cores e a percepção de detalhes finos. A fóvea é a região de maior acuidade visual.

Os impulsos elétricos são transmitidos dos fotorreceptores, passando pelas células horizontais, bipolares e amácrinas, até as células ganglionares que são de dois tipos: as células M e as células P, que se conectam às regiões periféricas e central da retina, respectivamente. Os longos axônios das células ganglionares formam os nervos ópticos, que por fim convergem no quiasma óptico. Nesta região em forma de X, os axônios provenientes das retinas nasais e temporais de cada olho são projetados contra e ipsilateralmente aos hemisférios cerebrais, respectivamente. A partir do quiasma óptico, uma pequena porção das células ganglionares é projetada ao colículo superior (aproximadamente 10% nos humanos), um antigo centro de processamento visual associado à coordenação de alguns movimentos oculares e reflexos visuais. O restante projeta-se às células do núcleo geniculado lateral (NGL) no tálamo.

A organização observada nas células ganglionares mantém-se no NGL, dando origem a duas vias com diferenças anátomo-funcionais: a divisão magnocelular e a divisão parvocelular. A primeira recebe sinais das células M e a segunda das células P. A divisão magnocelular tem células de resposta rápida e transitória, com alta sensibilidade ao baixo

contraste, cuja função está relacionada ao movimento e à percepção de profundidade. Por outro lado, a divisão parvocelular tem células de resposta lenta e sustentada, sensíveis ao alto contraste e sua função vincula-se à visão de cores e à acuidade.

Do NGL, via radiações ópticas, a informação é enviada ao lobo occipital onde está localizada a área visual primária (V1) ou córtex estriado. O processamento da informação aumenta em complexidade à medida que alcança outras áreas cerebrais, geralmente especializadas em processar um atributo ou característica da imagem vista. Na década de 80, a partir de estudos que realizaram ablações no cérebro de macacos, reconheceram-se duas grandes e específicas vias do processamento visual: a via ventral e a via dorsal (Mishkin, Ungerleider & Macko, 1983).

Na primeira delas, a partir da área visual primária, a informação percorre regiões ao longo do lobo temporal, e sua função está associada ao reconhecimento dos objetos. Em razão disso, a via ventral também é conhecida como via “o quê”. Por outro lado, no lobo parietal ocorre o processamento das informações relacionadas à localização espacial na via que ficou conhecida como dorsal (ou via “onde”).

Posteriormente, Goodale e Milner (1992) argumentaram que seria mais apropriado chamar a via dorsal de via “como”, e não mais “onde”, dada a sua repercussão não só na localização, mas também no planejamento e execução de ações. Há evidências de que estas vias recebem informações provenientes das vias parvo e magnocelular, respectivamente (Kessels, Postma & De Haan, 1999). No entanto, recentemente esta ideia foi questionada, ou seja, as células da via magnocelular não se projetam exclusivamente para a via dorsal (Callaway, 2005).

A via dorsal é mais curta e seus neurônios têm latências menores frente a estimulação visual, provavelmente por serem mais mielinizados e pelo seu papel na interação visuomotor, que requer respostas mais rápidas. Em contraste, a via ventral é mais longa e os neurônios possuem latências maiores, dado que para a identificação precisa de um objeto a velocidade não é tão fundamental. Cabe ressaltar, que ao expor que a velocidade de transmissão de informação na via ventral é mais lenta do que na via dorsal, faz-se referência a tempos inferiores a um décimo de segundo (entre 80 e 100 ms) para a chegada da informação no córtex temporal e de um tempo entre 100 e 130 ms para atingir o córtex pré-frontal (Rousselet et al., 2004b).

A partir de uma revisão de estudos neurofisiológicos, neuropsicológicos e psicológicos, Norman (2002) expôs o paralelo existente entre o funcionamento das vias dorsal e ventral e as abordagens ecológica e construtivista, respectivamente. Contudo, esta divisão entre as vias ventral e dorsal não significa que a informação presente em uma via não interaja ou não seja afetada pelas informações processadas na outra via mediante mecanismos de retroalimentação (*feedback*). Em um estudo neurofisiológico com macacos, encontrou-se que as informações processadas nas duas vias convergem para áreas superiores onde são integradas, por exemplo no córtex pré-frontal (Rao, Rainer & Miller, 1997). Nos últimos anos, alguns estudos questionaram tal divisão, mostrando que o processamento do reconhecimento e da localização de um objeto ocorre em paralelo e em ambas as vias (Konen & Kastner, 2008). Pode-se concluir que a separação entre via dorsal e ventral tem valor clínico e didático para compreender a complexidade do processamento visual.

Com relação às dicas de profundidade, as maiores evidências apontam que o processamento destas ocorra em áreas do córtex parietal, isto é, na via dorsal (“onde”). Esta afirmação é verdadeira tanto para o gradiente de textura (Tsutsui, Sakata & Naganuma, 2002) quanto para a perspectiva (Welchman, Deubelius, Conrad, Bülthoff & Kourtzi, 2005), embora existam argumentos favoráveis ao processamento desta na via ventral (Bruggemann, Yonas & Konczak, 2007). Independente desta controvérsia, poucas dúvidas pairam sobre a rapidez do processamento dos indícios de profundidade e seus efeitos sobre a percepção de tamanho.

Neste sentido, há o recente trabalho de Torro-Alves, Bacci, Bernardino e Fukusima (2012). A tarefa dos participantes consistia em comparar o tamanho de duas barras verticais apresentadas por apenas 100 ms sobre gradientes de textura gerados computacionalmente. Os resultados revelaram uma grande superestimação do tamanho (20%) dos estímulos inseridos no gradiente de perspectiva. No gradiente de compressão, este efeito mostrou-se bastante reduzido, com uma superestimação do tamanho de 8%. Ou seja, o gradiente de perspectiva teve um efeito 2 vezes maior sobre o tamanho percebido em comparação com o gradiente de compressão.

No entanto, o resultado mais interessante deste estudo advém da condição em que os estímulos estavam inseridos em um gradiente de textura formado pelas informações de perspectiva e compressão. Nesta situação, foi observada uma superestimação no tamanho dos estímulos de 28%, isto é, a soma dos efeitos individuais observados em cada gradiente. As distorções de tamanho descritas revelam que os gradientes de textura foram eficientemente

processados, pois afetaram o tamanho percebido. Assim, pode-se afirmar que tal processamento ocorreu em 100 ms ou menos e, além disso, este tempo foi suficiente para integrar duas informações de profundidade distintas.

Resultados semelhantes também foram encontrados em tarefas de discriminação de tamanho (Bernardino, 2008) e de busca visual (Champion & Warren, 2008). O primeiro estudo utilizou tempos breves de exposição (50, 100 e 200 ms), enquanto o segundo apresentou os estímulos com menor restrição temporal (100, 200, 400 e 800 ms). Apesar destas diferenças, ambos os estudos concluíram que a representação do tamanho levando-se em consideração a informação de profundidade ocorre em torno dos 100 ms.

Estudos anteriores com abordagem neurofisiológica também apontaram uma faixa de tempo semelhante para a integração das informações de tamanho e profundidade, oscilando entre 80 e 120 ms (Di Lollo, Enns & Resink, 2000; Tong, 2003). Em contraste, um estudo eletrofisiológico apontou que não somos sensíveis à perspectiva antes dos 90 ms e que a extração de informação de profundidade a partir da perspectiva ocorre num segundo nível de processamento, necessitando o envolvimento de áreas corticais superiores (Cauquil, Trotter & Taylor, 2006). Este resultado, no entanto, não invalida as evidências supracitadas sobre a velocidade de processamento das informações de tamanho e profundidade, na medida em que estabelece apenas um limite inferior antes do qual não haveria processamento, mas não indica um limite superior.

Ademais, Bernardino (2008) observou que as distorções de tamanho reduziram-se significativamente em apresentações de 200 ms. Uma hipótese para explicar este resultado é o recrutamento de mecanismos atentos, o que resultaria em menores distorções de tamanho à medida que aumenta o tempo de exposição e, conseqüentemente, o tempo disponível para processar as informações. Assim, a atenção como fator modulador da percepção de tamanho e profundidade é a discussão que empreendemos a seguir.

1.5. Atenção: fator modulador da percepção de tamanho e profundidade

A atenção é um processo que direciona recursos para o processamento mais completo de aspectos selecionados de um objeto ou imagem em detrimento dos não selecionados

(Palmer, 1999). Neste conceito estão contidas duas propriedades básicas da atenção: a capacidade e a seletividade, isto é, temos recursos limitados para processar as informações e assim é necessário selecionar as mais relevantes a cada momento. Goldstein (2009) apresenta uma interessante metáfora: frente à imensa quantidade de estímulos que nos chega de maneira contínua pelos órgãos dos sentidos, realizar o processamento de toda esta informação seria como tentar beber água de uma mangueira dos bombeiros. É simplesmente impossível.

A atenção, portanto, é um filtro que evita a sobrecarga do sistema nervoso, que tem uma capacidade limitada para processar as informações advindas do ambiente. Desta maneira, a informação proveniente de uma região atendida do campo visual é processada mais eficientemente do que as localizadas em uma área não atendida.

Há dois processos principais envolvidos no caráter direcional da atenção, que servem como base para sua divisão e classificação em atenção sustentada e atenção transitória, as quais foram caracterizadas por vários autores (por exemplo, Nakayama & Mackeben, 1989; Theeuwes, Kramer & Kingstone, 2004). A atenção sustentada (ou endógena) é o direcionamento voluntário desta para objetos ou partes de uma cena, em função de um objetivo definido. A atenção sustentada é um processo um pouco lento (300 ms) por relacionar-se à volição do indivíduo. Isto significa que ela está vinculada a mecanismos *top-down* ou de processamento nos níveis superiores da hierarquia cerebral. Pode-se estudá-la através de instruções específicas, que indicam aonde a atenção deve ser alocada pelo observador.

Em contraste, a atenção transitória (ou exógena) refere-se à alocação da atenção em um estímulo de maneira reflexiva, independente de nossos objetivos e/ou interesses. É um processo mais rápido (100 ms), ligado a mecanismos *bottom-up* ou aos estágios iniciais de processamento. Para a investigação da atenção transitória, em geral, um estímulo é apresentado repentinamente em alguma região do campo visual (dica ou *cue*), resultando na captura e direcionamento automático da atenção para esta região.

Os efeitos da atenção sobre a percepção visual já foram amplamente descritos em estudos anteriores; de maneira geral, revelam um incremento no desempenho em diversas tarefas perceptuais (para alguns exemplos, ver Cameron, Tai & Carrasco, 2000; Carrasco, Penpeci-Talgar & Eckstein, 2000; Spivey & Spirn, 2000; Yeshurun & Carrasco, 1999). Há algumas hipóteses para explicar esta melhora perceptiva gerada pela atenção. A primeira delas é a redução da incerteza de resposta, ou seja, a atenção afetaria a percepção ao nível de

decisão e não a representação sensorial dos estímulos. Um estudo cujos resultados se coadunam com esta hipótese é o de Prinzmetal, McCool e Park (2005). Uma interessante perspectiva para testar esta hipótese é o uso da Teoria de Detecção de Sinal (Green & Swets, 1988), que permite analisar os critérios de resposta dos participantes.

Outro mecanismo que pode ser o responsável por tal efeito é o aumento de sinal (*signal enhancement*), assim a atenção aumentaria e/ou facilitaria a representação sensorial do estímulo. Evidências para esta explicação foram encontradas por alguns autores, entre eles Carrasco, Ling e Read (2004), Ling e Carrasco (2006) e Lu e Doshier (1998). Por último, há a hipótese da redução dos distratores, isto é, a atenção diminuiria o efeito distrator dos estímulos não atendidos, ideia que encontra respaldo no trabalho de Doshier e Lu (2000).

Além disso, muitos estudos indicam um melhor desempenho dos participantes no campo visual inferior em diversas tarefas (para revisão, ver Qu, Song & Ding, 2006), fenômeno conhecido como assimetria do meridiano vertical. Segundo Previc (1990), esta assimetria teria sido forjada ao longo do processo evolutivo de nossa espécie em resposta às demandas ambientais. Nesta direção, também há evidências de que o campo visual inferior teria alta resolução espacial e, por isso, teria papel fundamental na representação do espaço próximo (peripessoal); enquanto o campo visual superior estaria relacionado ao processamento das informações mais distantes por sua grande resolução temporal (Zhou & King, 2002).

A hipótese prevalente para explicar esta assimetria é a alocação diferencial da atenção no campo visual inferior em detrimento do campo visual superior (He, Cavanagh & Intriligator, 1996). Esta hipótese encontra respaldo do ponto de vista neural, pois a porção de V1 que representa o campo visual inferior envia mais projeções para a via dorsal, sistema implicado na atenção, do que a parte correspondente ao campo visual superior (Gazzaniga & Ladavas, 1987). Entretanto, Talgar e Carrasco (2002) defendem que a assimetria do meridiano vertical ocorre devido a fatores visuais e não por alocação diferencial da atenção entre os campos visuais.

No que diz respeito ao efeito da atenção sobre a percepção de tamanho, a maioria dos estudos indica distorção no tamanho percebido devido à atenção, embora a direção desta distorção seja bastante controversa. Há estudos que descrevem que a atenção diminui o tamanho percebido de uma linha (ver Tsal, Shalev & Zakay, 2005), enquanto outros defendem exatamente o contrário (ver Prinzmetal & Wilson, 1997). O aumento no tamanho

percebido pela influência da atenção também foi encontrado por Anton-Erxleben, Henrich e Treue (2007). Estes autores apresentavam os estímulos precedidos por uma dica, em um paradigma de atenção transitória. Observaram-se superestimações no tamanho dos estímulos de até 12%, indicando que mecanismos *bottom-up* afetam a representação do tamanho, principalmente se os estímulos forem pequenos.

Ademais, a relação entre atenção e percepção de tamanho não é unilateral. Evidências recentes mostram que o tamanho é um atributo capaz de recrutar mecanismos *bottom-up* da atenção, mesmo que o estímulo em questão seja irrelevante para a tarefa (Proulx, 2010). Resultados semelhantes foram descritos por Proulx e Green (2011), porém amplia-se o escopo do estudo anterior. Os autores utilizaram uma versão modificada da ilusão de Müller-Lyer e encontraram que o estímulo maior foi detectado com mais eficiência. Note-se que, neste caso, a percepção baseia-se no tamanho aparente dos estímulos e não no tamanho físico.

A atenção também é alocada preferencialmente em regiões que possuem informações de profundidade, sejam elas provenientes de indícios pictóricos monoculares (Morita & Kumada, 2003) ou de disparidade binocular (He & Nakayama, 1995). Além disso, em um estudo que combinou medidas comportamentais e eletrofisiológicas, Parks e Corballis (2006) encontraram um menor tempo de reação e uma maior amplitude no componente sensorial P1 (100-160 ms) quando a atenção era dirigida a estímulos localizados mais distantes em relação aos observadores. Ambas as medidas mantiveram-se constantes para estímulos próximos. Os mesmos autores encontraram que mudanças atencivas de estímulos distantes para estímulos próximos são mais rápidas do que no sentido inverso.

Tomados em conjunto, estes resultados evidenciam que o processamento das informações de tamanho e de profundidade é eminentemente afetado por mecanismos *bottom-up* da atenção. Embora no clássico estudo de Posner (1980) tenha sido demonstrado que a atenção pode ser alocada a diferentes regiões do campo visual sem a necessidade de movimentos oculares, não se pode negligenciar a relação destes com a atenção para uma compreensão ampla deste tema. Afinal, movimentos oculares ocorrem amiúde em tempos superiores a 150 ms (Bourne 2006) e são precedidos pela alocação da atenção (Corbetta et al 1998; Hoffman & Subramaniam 1995, Rizzolatti, Riggio, Dascola & Umiltá, 1987). Há evidências recentes de que a relação entre atenção e movimentos sacádicos não é obrigatória e nem imediata, embora compartilhem circuitos anatômicos e sejam processos paralelos que funcionam em harmonia (para uma revisão, ver Awh, Armstrong & Moore, 2006).

1.6. Movimentos oculares: mecanismos *bottom-up* e *top-down*

Ao observar um ponto de fixação estático, o campo visual adota uma forma elíptica e compreende um ângulo visual de aproximadamente 160° no meridiano horizontal e 135° no meridiano vertical (Ruch & Fulton, 1960), embora o campo visual central não se estenda além dos 30° de ângulo visual. É importante ressaltar que as projeções que incidem sobre a fóvea subtendem um ângulo não maior do que 2° , correspondendo a área de maior acuidade visual e onde ocorre o processamento mais detalhado da informação visual. As imagens projetadas que subtendem ângulos entre 2° e 10° incidem sobre uma região chamada parafoveal e acima de 10° é considerado que o objeto está representado na periferia da retina.

Os olhos dos humanos movimentam-se independentemente da cabeça ou do corpo, graças a seis pares de músculos oculomotores, dois responsáveis pelo movimento horizontal (reto médio e lateral) e quatro vinculados ao movimento vertical (reto superior e inferior e oblíquo superior e inferior). Esta capacidade confere grande vantagem adaptativa ao permitir que os olhos sejam continuamente posicionados sobre as regiões mais salientes e informativas, de modo que estas incidam sobre a fóvea para o processamento mais eficiente das informações ali contidas, gerando uma percepção mais acurada do espaço (Leigh & Zee, 1991).

A partir desta constatação, revela-se uma das propriedades básicas dos movimentos oculares: os olhos têm um comportamento propositivo e não se movem aleatoriamente de uma posição a outra. Outra característica básica refere-se ao fato de que os olhos nunca estão totalmente estáticos. A retina adapta-se rapidamente e a percepção desaparece em menos de 100 ms (Riggs, Ratiliff, Cornsweet & Cornsweet, 1953). Mesmo que nos sejam imperceptíveis, há a ocorrência de movimentos oculares miniatura, que são involuntários e de amplitude extremamente pequena, a despeito da impressão que nosso olhar encontra-se fixo sobre uma posição.

Além deste, há outros tipos de movimentos oculares, que segundo Cabestrero, Conde-Guzón, Crespo, Grzib e Quirós (2005), podem ser assim classificados: 1) Movimentos compensatórios, que são movimentos reflexos que dependem da atuação do sistema vestibulo-ocular, cuja função é estabilizar a posição dos olhos em relação ao ambiente frente o movimento da cabeça ou do corpo; 2) Movimentos de perseguição, que são movimentos de

acompanhamento de objetos que se deslocam lentamente, caracterizam-se por serem suaves e automáticos em geral; 3) Movimentos de vergência, que são movimentos coordenados de ambos os olhos, movendo-os em direções opostas no meridiano horizontal para focalizar um objeto; e 4) Movimentos sacádicos, que são movimentos rápidos, realizados 3 vezes por segundo em média, tendo um papel fundamental na exploração do ambiente, como será descrito a seguir.

Os movimentos sacádicos ou sacadas são classificados como balísticos, dada a sua velocidade e a dificuldade em interrompê-los uma vez iniciados. Realizamos, em média, entre 2 e 4 sacadas por segundo. Cabe ressaltar que durante estes movimentos, a apreensão da informação ambiental é quase completamente interrompida, fenômeno esse chamado de supressão sacádica. Além disso, eles possuem uma finalidade definida, que é a de dispor os objetos ou partes da cena na região da fóvea, o que permite o processamento mais detalhado da informação nela disposta (Henderson & Hollingworth, 1998).

Entre os movimentos sacádicos são observadas as fixações, momentos de relativa estabilidade dos olhos sobre uma região. Diz-se “relativa”, pois como exposto anteriormente, os olhos não permanecem totalmente estáticos a fim de evitar a adaptação da retina e o consequente desaparecimento da imagem. Desta maneira, para determinar o que é uma fixação, geralmente estabelece-se dois critérios, um espacial e outro temporal. Uma definição amplamente aceita é a proposta por Salvucci e Goldberg (2000), na qual é considerada uma fixação caso o olhar do observador não se desloque por uma área quadrada maior que 1° de ângulo visual por um período de 100 ms.

Os movimentos oculares são registrados através de diversas técnicas, que se desenvolveram e aumentaram em sofisticação nos últimos anos, permitindo registros com resolução temporal menor e resolução espacial mais precisa (para uma revisão, ver Cabestrero et al., 2005). Apesar da crescente tecnologia e da grande quantidade de pesquisas investigando os movimentos oculares, não se esgotaram as possibilidades de investigação. Há aspectos ainda desconhecidos ou controversos relacionados à aquisição, à representação e ao processamento dinâmico das informações provenientes do ambiente.

Os movimentos sacádicos guardam uma estreita relação com a atenção, dado que ambos compartilham vias neurais (Awh et al., 2006) e servem a uma mesma finalidade: selecionar a informação mais relevante do meio e direcionar e recrutar os recursos sensoriais e cognitivos para o processamento mais acurado e completo da informação-alvo. Cabe salientar

que as sacadas podem ser involuntárias ou voluntárias, estando relacionadas assim à dinâmica temporal da percepção. Isto quer dizer que em um primeiro estágio do processamento, onde prevalecem os mecanismos atentos *bottom-up*, as sacadas seriam geradas involuntariamente. Por outro lado, os processos de ordem superior dariam lugar aos mecanismos *top-down* e as sacadas estariam sob maior controle volitivo.

Os movimentos sacádicos podem ser estudados, por exemplo, através de sua amplitude. Este parâmetro revela a extensão com que o campo visual é explorado. Quanto maiores as amplitudes sacádicas, menores são as exigências para o processamento da informação contida numa região (Miura, 1990). A duração de uma sacada e sua velocidade também são parâmetros a serem investigados, embora estejam fortemente correlacionados à amplitude: maior amplitude do movimento sacádico indica maior duração e menor velocidade do mesmo.

A direção das sacadas é mais um parâmetro que nos fornece importantes indícios sobre o processamento das informações visuais. Foulsham e Kingstone (2010) encontraram uma maior probabilidade de ocorrência de movimentos sacádicos no eixo vertical do que no eixo horizontal. Além disso, os resultados mostram uma prevalência de sacadas dirigidas à região inferior das imagens. De modo muito interessante, esta prevalência manteve-se ao se apresentar as imagens invertidas (rotação de 180°). Neste caso, a maioria das sacadas foi dirigida para cima, onde estava localizada a porção inferior da imagem. Estes resultados se coadunam com a vantagem no processamento de objetos localizados no campo visual inferior descrita em outros estudos (por exemplo, Portin, Vanni, Virsu & Hari, 1999). Esta preferência também é observada desde a primeira sacada, cuja direção deve ser pré-programada ou baseada nos estágios iniciais de processamento (Henderson, Weeks & Hollingworth, 1999), isto é, está relacionada aos mecanismos *bottom-up*.

A introdução de informações de profundidade também afeta propriedades dos movimentos oculares. Em um elegante estudo, Jansen, Onat e Koni (2009) empreenderam uma importante distinção entre as sacadas dirigidas ao campo visual inferior e as sacadas dirigidas para objetos percebidos como próximos ao observador. Eles utilizaram imagens naturais em 2D e em 3D e concluíram que, na verdade, as sacadas não são preferencialmente dirigidas para baixo e sim para objetos próximos. Em contraste, Wexler e Quartí (2008) indicaram que ao menos as sacadas iniciais são dirigidas às regiões de maior profundidade, principalmente a primeira sacada.

Assim como os movimentos sacádicos, o estudo da dimensão espacial e temporal das fixações é fundamental para elucidar o envolvimento dos mecanismos *bottom-up* e *top-down* na representação do espaço. A fixação é um processo complexo, no qual é analisado a região ou o objeto disposto na fóvea. Após a aquisição das informações necessárias para o processamento dessa região ou objeto, a fixação é interrompida e dá lugar a um movimento sacádico (Hayhoe & Ballard, 2005).

A duração de uma fixação oscila entre 100 e 1000 ms, com um valor médio de aproximadamente 300 ms. Henderson e Hollingworth (1998) encontraram que as durações das fixações eram muito similares, independentemente se as imagens observadas eram representações tridimensionais, cenários fotografados ou suas contrapartes reproduzidas pictoricamente. Segundo Henderson (1992), durante uma fixação podem ser identificadas três fases distintas: 1) processamento da informação foveal, 2) realização de uma amostra da periferia visual, e 3) planejamento do próximo movimento sacádico.

A duração da primeira fase é de aproximadamente 100 ms (Kirchner & Thorpe, 2006; Viviani, 1990) e o restante do tempo é utilizado para a realização das duas fases seguintes. Neste sentido, Van Diepen, De Graef e d'Ydewalle (1995) defendem que a primeira fixação seria mais longa para o processamento inicial da imagem, o que repercutiria também sobre a programação sacádica e exploração subsequentes. Por outro lado, há evidências de que as primeiras fixações são mais rápidas e o processamento mais eficiente da imagem ocorre posteriormente, o que é evidenciado pelo aumento da duração das fixações no decorrer da observação das imagens (Antes, 1974; Jansen et al., 2009). Apesar destas evidências de que a duração de uma fixação reflete o processamento das informações, Vaughan (1982) defende que a duração é pré-programada e não guarda relação com o tempo de processamento visual.

Desde o clássico estudo de Yarbus em 1967 (para uma interessante revisão, ver Tatler, Wade, Kwan, Finlady & Velichkovsky, 2010) e até antes (Buswell, 1935), reconhece-se que as áreas mais informativas de uma imagem recebem mais fixações e são observadas por mais tempo do que as áreas menos informativas. É importante salientar que a definição do que é uma zona informativa pode aludir tanto a mecanismos *bottom-up* quanto a mecanismos *top-down*.

No primeiro caso, há evidências de que as propriedades visuais da cena afetam o controle temporal e espacial das fixações. Parkhurst e Niebur (2003) encontraram um padrão oposto na localização das fixações, dependendo do conteúdo da imagem observada: maior

concentração de fixações no campo visual inferior para cenas urbanas e no campo visual superior para cenas de ambientes internos. Ademais, a grande consistência e pequena variabilidade entre os participantes para selecionar a localização das primeiras fixações é mais um argumento a favor da mediação *bottom-up* para este processo (Parkhurst, Law & Niebur, 2002; Tatler, Baddeley & Gilchrist, 2005).

Desta maneira, a ideia de mapas de saliência é consonante com este argumento (ver Henderson & Hollingworth, 1998). As diferentes regiões de uma imagem receberiam pesos, indicando a probabilidade das regiões serem alvo de fixações. Estes pesos são determinados de acordo com as características perceptivas (cor, contraste, textura, etc.) da região em questão. Os mapas de saliência seriam contínua e dinamicamente modificados à medida que a imagem é explorada.

Por outro lado, Loftus e Mackworth (1978) mostraram que determinar se uma área é informativa ou não depende da natureza semântica da informação, o que está relacionado a processos cognitivos ou *top-down*. Os autores apresentavam uma cena (por exemplo, uma fazenda) e inseriam objetos congruentes (trator) ou incongruentes (polvo). A análise dos parâmetros oculares indicou que os objetos incongruentes recebiam mais fixações e eram inspecionados por mais tempo.

Nos últimos anos, surgiram várias evidências de que o processo de seleção da posição espacial das fixações é mediado fortemente por mecanismos *top-down* em detrimento dos processos *bottom-up* (por exemplo, Cristino & Baddeley, 2009). Neste sentido, as instruções dadas aos participantes (Watson, Brennan, Kingstone & Enns, 2010) e a tarefa realizada afetam os parâmetros oculares investigados (Castelhano, Mack e Henderson, 2009). Além disso, observou-se que as fixações tornam-se mais longas à medida que aumenta a dificuldade em tarefas de discriminação (Hooge & Erkelens, 1996). Consonante com estes resultados, também há indícios de que os neurônios da área V1 disparam de uma maneira tarefa-dependente, de modo a extrair e processar as informações que serão úteis naquele contexto (Gilbert, Sigman & Crist, 2001).

Em uma posição mais conciliadora entre os mecanismos *bottom-up* e *top-down* encontra-se o estudo Tatler (2007). Seus resultados sugerem que as características da imagem e o tipo de tarefa interagem e influenciam na determinação da posição espacial das fixações ao longo da observação de uma cena. Assim, não existiria uma dicotomia tão severa entre estes mecanismos.

Ademais, os estudos sobre percepção de profundidade e tamanho valendo-se do registro dos movimentos oculares ainda são poucos. Por exemplo, Jansen et al. (2009) encontraram que a adição de informações de profundidade alterou os parâmetros oculares. Dentre os resultados mais significativos, pode-se citar que a introdução de informações de profundidade, principalmente a disparidade binocular, levou a um maior número de fixações e maior distribuição espacial destas. Com relação à percepção de tamanho, há um interessante estudo com o peixe (*Toxotes chatareus*), conhecida no Brasil como peixe arqueiro sete-manhas (Segev, Schneidman, Goodhouse & Berry, 2007). Os autores treinaram os animais para realizar discriminações de tamanho e registraram os movimentos oculares e a atividade neuronal de células ganglionares da retina. Os resultados indicaram que o início da fixação é o momento crítico para extrair as informações de tamanho.

Enfim, as evidências mostram que a percepção de tamanho e a percepção de profundidade são forjadas em tempos curtos, porém a dinâmica temporal destes processos ainda não foi completamente elucidada. Tampouco se compreende de maneira ampla o envolvimento dos mecanismos *bottom-up* e *top-down* nestes processos.

Deste modo, foram realizados dois experimentos com o objetivo de investigar a dinâmica temporal da percepção de tamanho de objetos inseridos em gradientes de textura, provenientes de desenhos gerados computacionalmente ou de fotografias com elementos naturais. Esperava-se, portanto, alcançar uma maior compreensão sobre os processos perceptuais e cognitivos envolvidos no processamento em conjunto das informações de tamanho e de profundidade.

No primeiro experimento, utilizou-se um método psicofísico e analisaram-se os componentes sensoriais (distorção perceptual e a sensibilidade discriminativa) observados em uma tarefa de comparar tamanho relativo. No segundo experimento, registraram-se os movimentos oculares dos participantes em uma tarefa semelhante. Verificou-se assim, o envolvimento dos mecanismos *bottom-up* e *top-down* no processamento das informações de tamanho e profundidade em condições com alta e baixa restrição temporal.

2. Objetivos gerais

O presente estudo teve por objetivo investigar se diferentes tempos breves de exposição afetam o julgamento de tamanho de estímulos inseridos em gradientes de textura, com informações de profundidade provenientes de desenho gerado computacionalmente ou de fotografia com elementos naturais. Além disso, teve-se por objetivo investigar a ocorrência de movimentos oculares nas comparações de tamanho realizadas em tempos breves de exposição, bem como identificar o padrão dos movimentos oculares em uma condição com menor restrição temporal.

3. Experimento 1

3.1 Objetivo

Investigar como tempos breves de exposição dos gradientes de textura (desenho gerado computacionalmente e fotografia com elementos naturais) afetam a comparação de tamanho de estímulos neles inseridos.

3.2 Método

3.2.1 Participantes

Participaram deste experimento 96 voluntários distribuídos aleatoriamente em 12 grupos equitativos, cada um com 3 homens e 5 mulheres. Os voluntários eram alunos da Universidade de São Paulo de Ribeirão Preto, com idade entre 18 e 35 anos, com acuidade visual normal ou corrigida em ambos os olhos. Antes de submeterem-se ao experimento, todos os voluntários assinaram o Termo de Consentimento Livre e Esclarecido (ANEXO A), aprovado pelo Comitê de Ética em pesquisa da FFCLRP-USP (ANEXO B), conforme as normas vigentes no Brasil sobre experimentos com humanos.

3.2.2. Material e equipamentos

O experimento foi realizado por meio de um computador Pentium IV, 992 MB RAM, com placa de vídeo NVIDIA GeForce 8400GS, plataforma Windows XP e um monitor LG Flatron ez T910B Plus (19”), sob resolução de 1024 x 768 pixels e frequência de atualização de 85 Hz. Foi utilizado um aparelho Ortho-Rather da Bausch & Lomb Optical Co. para medir a acuidade visual dos participantes.

Um programa foi desenvolvido em Lazarus – Free Pascal 0.9.28.2 para gerar e apresentar os estímulos e coletar as respostas dos participantes durante a execução do experimento em uma sala escura. Um apoio para o queixo foi utilizado para posicionar adequadamente a cabeça dos participantes em frente ao monitor a uma distância de 57 cm e

uma venda para que se mantivessem sob condição monocular com o olho direito (olho esquerdo vendado) durante a realização do experimento.

Os estímulos eram dois círculos negros (ambos com luminância de $1,1 \text{ cd/m}^2$) apresentados simultaneamente: o padrão, apresentado no campo visual inferior com tamanho de 100 *pixels* ou $3,4^\circ$ de ângulo visual de diâmetro; e outro o teste, apresentado no campo visual superior cujo tamanho variou de $2,89^\circ$ a $3,91^\circ$ de ângulo visual de diâmetro, com diferença de $0,102^\circ$ entre eles, totalizando 11 possíveis estímulos teste. Estes estímulos foram apresentados sobre três fundos de tela, como pode ser observado na Figura 1: sem textura (1a); gradiente de textura em perspectiva gerado computacionalmente (1b); e gradiente de textura fotografado com elementos naturais (1c). Com o objetivo de tornar o texto mais claro e direto, adotou-se a seguinte nomenclatura ao longo da tese para os fundos de tela: controle (1a), gradiente de perspectiva (1b) e gradiente fotografado (1c). Estas telas possuíam luminância média⁴ de 61 cd/m^2 , 60 cd/m^2 e 57 cd/m^2 , respectivamente. Para atingir este valor de luminância, o gradiente fotografado foi manipulado com o auxílio do programa CorelDRAW X3.

Figura 1 - Exemplos de estímulos na condição controle (a), na condição com gradiente de perspectiva (b) e na condição com gradiente fotografado (c). Em todos os exemplos, ambas os círculos têm o mesmo tamanho ($3,4^\circ$ de ângulo visual no tamanho original).

Cada um dos participantes realizou as comparações de tamanho em um único tempo de exposição dos estímulos (50, 100, 150 ou 200 ms) e um único fundo de tela (controle ou gradiente de perspectiva ou gradiente fotografado). A combinação de um tempo de exposição, dentre quatro possíveis, e um fundo de tela, dentre três possíveis, resultou em 12 grupos experimentais independentes, aos quais os participantes foram designados aleatoriamente.

⁴ O cálculo da luminância média foi realizado a partir da medida de luminância em 9 regiões diferentes do monitor. Esta variação da luminância em função do fundo de tela está dentro de limites esperados ($6,2 - 7,5 \text{ cd/m}^2$) e não afeta os resultados encontrados. Para mais detalhes, ver Santos, Nogueira e Simas (2005).

3.2.3 Procedimento

Antes de iniciar o experimento, cada participante teve sua acuidade visual avaliada pelo aparelho Ortho-Rater da Bausch & Lomb Optical Co. Os participantes com acuidade normal ou corrigida em ambos os olhos (6/6 ou superior) participaram do experimento. Nenhum participante foi excluído por esse critério. Ao cumprir esse requisito, o Termo de Consentimento Livre e Esclarecido era entregue ao participante para a leitura e, caso concordasse com as informações nele contidas, solicitava-se que o assinasse.

Cada participante aceito foi designado aleatoriamente a um dos grupos. Em sessão individual, o participante sentou-se em frente ao monitor e posicionou sua cabeça no apoiador de queixo que se encontrava a 57 cm de distância do centro do monitor. As instruções eram apresentadas na tela do computador (ANEXO C) e lidas pelo pesquisador para o participante, de modo que este não ficasse com nenhuma dúvida. Após a colocação da venda no olho esquerdo do participante, era iniciada a primeira das duas sessões experimentais. O intervalo entre as sessões foi determinado pelo próprio participante.

Cada tentativa consistiu da exposição prévia de um ponto de fixação no centro da tela por 500 ms, seguido pela apresentação dos pares de estímulos sobre um único fundo de tela (controle ou gradiente de perspectiva ou gradiente fotografado) e um mesmo tempo de exposição (50 ou 100 ou 150 ou 200 ms), dependendo do grupo para o qual o participante foi designado. Depois uma tela em branco era inserida durante 500 ms e, em seguida, apresentou-se uma tela para a coleta da resposta. Após a resposta ser dada, uma tela aparecia e o próximo estímulo era apresentado somente após o participante pressionar a tecla de espaço. Isso permitia que o participante tivesse o controle sobre a apresentação dos estímulos e pudesse, caso necessário, descansar entre as tentativas.

Os participantes eram instruídos a indicar pelo teclado do computador, em cada apresentação de estímulos, qual dentre os dois círculos apresentados era o maior: o que apareceu no campo visual superior ou o que apareceu no campo visual inferior. Antes de iniciar o experimento, o participante realizou um treino com 10 tentativas para se familiarizar com a tarefa.

Os estímulos foram apresentados pelo método psicofísico das escadas duplas (Cornsweet, 1962). Houve, portanto, uma série ascendente em que o tamanho inicial do estímulo teste era menor (88 *pixels*) do que o do estímulo padrão e uma série descendente na qual o tamanho do estímulo teste inicial era maior (112 *pixels*) do que o do estímulo padrão.

Para evitar que os participantes previssem a qual série os estímulos pertenciam, as apresentações destes em ambas as séries foram intercaladas aleatoriamente.

Esse método prevê que a intensidade do próximo estímulo a ser apresentado não é predeterminada, mas sim calculada durante o experimento, baseada na história prévia dos estímulos expostos e das respostas emitidas. Esse cômputo pode ser realizado através de diversas regras e, dependendo da escolhida, um ponto específico da função psicométrica será estimado (Levitt, 1971).

Nesse experimento foram utilizadas duas regras distintas e, por esse motivo, foram realizadas duas sessões experimentais. Em uma sessão, a regra para o ajuste do tamanho do estímulo na apresentação seguinte foi *1up-1down*. Isso significa que o tamanho do estímulo teste foi diminuído em um passo (fixo em 3 *pixels*), se na tentativa anterior ele foi julgado maior; ou aumentado um passo, caso tenha sido julgado como menor. Essa regra permite a estimativa do ponto de 50% ou Ponto de Igualdade Subjetiva (PIS) na função psicométrica, que é a intensidade do estímulo teste que na média é estimado como igual ao estímulo padrão.

A segunda regra utilizada foi a *1up-2down*, a qual define que um passo foi aumentado após uma resposta “menor” e que um passo foi diminuído após duas respostas “maior”. Esta regra estima o ponto de 70,7% na função psicométrica, indicando a intensidade do estímulo teste em que a probabilidade de uma sequência de respostas “menor” é igual à de uma sequência de respostas “maior”. Para atingir as 24 reversões em cada sessão foram necessárias cerca de 100 tentativas, o que demorava em torno de 10 minutos.

3.2.4 Análise de dados

Em cada um dos grupos, para as comparações de tamanho realizadas pelos participantes na sessão com a regra *1up-1down*, foi calculado o PIS, que é a intensidade estimada do estímulo perceptivelmente igual ao do estímulo padrão. Este foi calculado através da média dos tamanhos dos estímulos teste nas últimas 18 tentativas da sessão em que houve reversões. O PIS abaixo de 100 *pixels*, tamanho do estímulo padrão, indica que o estímulo teste foi perceptivelmente superestimado, ou seja, ajustado menor que o estímulo padrão; enquanto um PIS maior do que 100 *pixels* indica que o estímulo teste foi perceptivelmente subestimado, ajustado maior que o padrão.

Na sessão com a regra *1up-2down*, calculou-se também a média do tamanho do estímulo teste nas últimas 18 reversões, o que correspondia aproximadamente a 70,7% das

respostas em que os participantes julgaram o estímulo teste maior do que o estímulo padrão. Com esses dois pontos da função psicométrica (50% ou PIS e 70,7%) foi possível estimar, através do ajuste destes à função cumulativa normal, a inclinação da função psicométrica. Esta é um indicador diretamente proporcional à sensibilidade do sistema visual para realizar determinada tarefa. Isto é, a medida da inclinação informa sobre a alteração do desempenho dos observadores causada pela intensidade dos estímulos.

As inclinações das funções psicométricas e os PIS foram submetidos a uma análise de variância (ANOVA) com dois fatores independentes (“fundo de tela” e “tempo de exposição”). Estas ANOVAs foram realizadas após verificarem-se os requisitos de normalidade e de homocedasticidade das variâncias através dos testes de Shapiro-Wilk e de Levene, respectivamente. Para identificar os níveis dos fatores responsáveis por diferenças significativas foi utilizado o pós-teste de Bonferroni. A análise estatística foi realizada com o auxílio do programa SPSS 17.0.

3.3 Resultados

3.3.1 Inclinação da Função Psicométrica

Um parâmetro importante para se analisar as comparações de tamanhos é a inclinação da função psicométrica, a qual oferece informações sobre a sensibilidade do participante em discriminar diferenças entre os estímulos. As médias e os respectivos erros padrão dessas inclinações para cada condição experimental de comparação de tamanhos são mostrados na Figura 2. Os dados foram submetidos a uma ANOVA entre participantes de dois fatores independentes, com o seguinte modelo: 3 (fundos de tela: controle, gradiente de perspectiva e gradiente fotografado) x 4 (tempos de exposição: 50, 100, 150 e 200 ms). A ANOVA mostrou efeitos significativos para o fator “tempo de exposição” ($F_{3,84} = 5,641$, $p = 0,001$), porém o fator “fundo de tela” ($F_{2,84} = 0,144$, $p = 0,866$) e a interação entre os fatores não foram significativos ($F_{6,84} = 0,633$, $p = 0,704$). O pós-teste de Bonferroni ($p = 0,001$) aplicado aos tempos de exposição revelou que os participantes tiveram uma maior sensibilidade para discriminar a diferença de tamanho entre os estímulos aos 200 ms em comparação aos julgamentos realizados aos 50 ms.

Figura 2 - Inclinação média das funções psicométricas para comparações de tamanho realizadas nas 3 condições de fundo de tela (controle, gradiente de perspectiva e gradiente fotografado) e nos 4 tempos de exposição investigados (50, 100, 150 e 200 ms). As barras de erro referem-se ao erro padrão da média.

* $p = 0,001$, inclinação menor aos 50 ms do que aos 200 ms.

De maneira geral, nota-se pelas inclinações médias das funções psicométricas, que os participantes discriminaram melhor o tamanho dos estímulos à medida que o tempo de exposição aumentava. Este efeito ocorreu independente do gradiente de textura, indicando que estes não afetaram de maneira diferente a sensibilidade do sistema visual em uma tarefa de comparar tamanhos em apresentações breves.

3.3.2 Ponto de Igualdade Subjetiva (PIS)

Caso os fundos de tela e os tempos de exposição investigados não afetassem as comparações de tamanho, seria esperado um PIS com valor igual ao estímulo padrão (100 *pixels*). Na Figura 3 são apresentados os valores médios de PIS com os respectivos erros padrões da média para cada um dos 12 grupos experimentais. Nela também é mostrada uma linha negra horizontal no valor do estímulo padrão (100 *pixels*).

O estímulo teste sempre foi apresentado no campo visual superior, portanto valores abaixo de 100 *pixels* revelam uma superestimação do estímulo teste em relação ao estímulo padrão e valores acima deste valor indicam subestimação. Este último caso ocorreu em apenas um dos 12 grupos experimentais (gradiente fotografado - 100 ms) e, ainda assim, de maneira sutil (PIS = 100,62 *pixels*). Cabe ainda ressaltar que a partir da Figura 3 depreende-se uma

importante observação: a superestimação do estímulo teste foi modulada diferencialmente de acordo com o fundo de tela e o tempo de apresentação dos estímulos.

Figura 3 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado) e nos 4 tempos de exposição (50, 100, 150 e 200 ms). As barras de erro referem-se ao erro padrão da média. A linha negra horizontal indica o tamanho do estímulo padrão (100 pixels).

Para investigar se as diferenças observadas na Figura 3 eram significativas realizou-se uma ANOVA entre participantes e de dois fatores independentes, com o seguinte modelo: 3 (fundos de tela: controle, gradiente de perspectiva e gradiente fotografado) x 4 (tempos de exposição: 50, 100, 150 e 200 ms). Para o fator “tempo de exposição” não foi encontrada diferença significativa ($F_{3,84} = 1,681$, $p = 0,177$). Por outro lado, a ANOVA revelou diferenças significativas para o fator “fundo de tela” ($F_{2,84} = 4,355$, $p = 0,016$) e para a interação entre os fatores “tempo de exposição” x “fundo de tela” ($F_{6,84} = 3,782$, $p = 0,002$).

Para verificar a natureza desta interação entre “tempo de exposição” e “fundo de tela”, isto é, compreender como o tempo de exposição modulou as comparações de tamanho em cada um dos fundos investigados, realizou-se três ANOVAs com um fator independente (“tempo de exposição”). Esta análise foi realizada com o método recomendado por Howell

(2002), no qual os valores de F são recalculados, utilizando o erro quadrado médio da análise anterior no lugar dos calculados para os efeitos simples.

Para a condição experimental sem gradiente de textura (controle) não foi encontrada diferença entre os PIS, de acordo com o tempo de exposição dos estímulos ($F_{3,28} = 0,186$, $p = 0,905$). Observe na Figura 4 que os valores de PIS foram muito semelhantes, houve uma pequena superestimação em todos os tempos de exposição investigados. Provavelmente este resultado foi devido à altura no campo visual, a qual pode ser utilizada como uma informação de profundidade, afetando assim as comparações de tamanho (Epstein, 1966; Tozawa, 2010).

Figura 4 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas na condição controle nos 4 tempos de exposição (50, 100, 150 e 200 ms). As barras de erro referem-se ao erro padrão da média.

Para a condição com gradiente de perspectiva foram observadas diferenças significativas entre os tempos de exposição ($F_{3,28} = 3,405$, $p = 0,021$). O pós-teste de Bonferroni ($p = 0,015$) revelou uma grande superestimação do tamanho aos 100 ms ($\bar{x} = 94,59$, $EPM = 0,82$) em comparação aos julgamentos realizados aos 200 ms ($\bar{x} = 99,25$, $EPM = 0,85$). Entre o tempo de 200 ms e o de 150 ms ($\bar{x} = 95,56$, $EPM = 1,14$) foi observado um efeito marginal ($p = 0,083$). Observa-se na Figura 5, que o valor do PIS está próximo de 100 pixels no tempo de 50 ms, indicando poucos desvios na resposta. No tempo de 100 ms ocorre uma grande superestimação do estímulo teste, sugerindo que o gradiente em questão

foi processado pelo sistema visual dos participantes. Nos tempos subsequentes, os desvios diminuem e os resultados vão em direção à constância do tamanho.

Figura 5 - Pontos de Igualdade Subjetiva (PIS) médios para comparações de tamanho realizadas no gradiente de perspectiva nos 4 tempos de exposição (50, 100, 150 e 200 ms). As barras de erro referem-se ao erro padrão da média.

* $p = 0,015$, superestimação do estímulo teste é maior no tempo de exposição de 100 ms em comparação com o de 200 ms.

Nas comparações de tamanho realizadas no gradiente fotografado (Figura 6), mais uma vez os tempos de apresentação modularam diferencialmente os tamanhos percebidos ($F_{3,28} = 5,6532$, $p = 0,001$). O pós-teste de Bonferroni indicou que os PIS são diferentes ($p = 0,007$) entre os tempos de 150 ms ($\bar{x} = 94,30$, EPM = 1,37) e o de 100 ms ($\bar{x} = 100,63$, EPM = 0,80).

Figura 6 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas no gradiente fotografado nos 4 tempos de exposição investigados (50, 100, 150 e 200 ms). As barras de erro referem-se ao erro padrão da média.

* $p = 0,007$, superestimação do estímulo teste é maior no tempo de exposição de 150 ms em comparação com o de 100 ms.

Nesta condição, os resultados indicam pouco desvio nas comparações realizadas nos tempos de exposição de 50 ms e 100 ms, seguido por uma grande superestimação aos 150 ms. Em seguida, no tempo de 200 ms, os desvios voltam a diminuir, sugerindo que o PIS em tempos maiores alcançaria valores próximos ao de 100 *pixels* (tamanho do estímulo padrão).

Ademais, para dar prosseguimento à investigação da interação observada, os dados foram submetidos a 4 ANOVAs com um fator independente (“fundo de tela”), mais uma vez utilizado o método sugerido por Howell (2002). Nas comparações de tamanho realizadas aos 50 ms, os fundos de tela não modularam diferencialmente a percepção de tamanho ($F_{2,21} = 0,278$, $p = 0,760$). Na Figura 7, observa-se que as distorções de tamanho foram muito semelhantes.

Figura 7 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas aos 50 ms nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado). As barras de erro referem-se ao erro padrão da média.

Por outro lado, os fundos de tela afetaram o tamanho percebido nas comparações realizadas aos 100 ms ($F_{2,21} = 10,444$, $p = 0,001$). O pós-teste indicou que neste tempo de exposição, o estímulo teste foi fortemente superestimado na condição com o gradiente de perspectiva ($\bar{x} = 94,59$, $EPM = 0,82$) em comparação com os fundos de tela controle ($\bar{x} = 98,43$, $EPM = 1,17$) e gradiente fotografado ($\bar{x} = 100,62$, $EPM = 0,80$), como pode ser observado na Figura 8.

Figura 8 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas aos 100 ms nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado). As barras de erro referem-se ao erro padrão da média.

* $p = 0,027$, superestimação do estímulo teste é maior no gradiente de perspectiva em comparação com o controle.

$p = 0,001$, superestimação do estímulo teste é maior no gradiente de perspectiva em comparação com o gradiente fotografado.

No tempo de exposição de 150 ms, a ANOVA também revelou diferenças significativas nas comparações de tamanho, de acordo com o fundo de tela sob o qual os estímulos eram apresentados ($F_{2,21} = 4,138$, $p = 0,031$). O pós-teste revelou que as distorções de tamanho foram maiores no gradiente fotografado ($\bar{x} = 94,30$, $EPM = 1,37$) em comparação com o controle ($\bar{x} = 99,48$, $EPM = 1,46$). O PIS médio abaixo de 100 *pixels* indica que a distorção supracitada ocorreu no sentido de superestimar o estímulo teste (Figura 9).

Figura 9 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas aos 150 ms nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado). As barras de erro referem-se ao erro padrão da média.

* $p = 0,031$, superestimação do estímulo teste é maior no gradiente fotografado em comparação com o controle.

Por fim, o padrão das distorções de tamanho aos 200 ms foi semelhante ao observado aos 150 ms. Na Figura 10, nota-se que a maior superestimação do estímulo teste ocorreu no gradiente fotografado. Além disso, as distorções de tamanho tornam-se menores nas comparações realizadas no gradiente de perspectiva e, diminuem ainda mais, na condição controle. No entanto, a ANOVA não indicou que estas diferenças eram significativas ($F_{2,21} = 2,598$, $p = 0,098$).

Figura 10 - Ponto de Igualdade Subjetiva (PIS) médio para comparações de tamanho realizadas aos 200 ms nos 3 fundos de tela (controle, gradiente de perspectiva, gradiente fotografado). As barras de erro referem-se ao erro padrão da média.

3.4 Discussão

O presente experimento teve por objetivo verificar a dinâmica temporal da percepção de tamanho e de profundidade, principalmente nos estágios iniciais de processamento. Realizou-se, para isso, uma tarefa psicofísica de comparação de tamanho de estímulos inseridos em gradiente de textura gerado computacionalmente (perspectiva) e gradiente de textura fotografado com elementos naturais. Os estímulos eram dois círculos apresentados no meridiano vertical por tempos muito breves. Analisou-se a inclinação da função psicométrica como índice da sensibilidade discriminativa dos participantes e o ponto de igualdade subjetiva para indicar as distorções de tamanho e, consequentemente, o processamento subjacente e a integração das informações de tamanho e profundidade. Desta maneira, maiores inclinações revelam uma melhor capacidade de discriminação dos estímulos e maiores distorções de tamanho indicam o efeito dos gradientes de textura e do tempo de exposição sobre o processamento de tamanho e profundidade.

Os resultados mostraram que a inclinação da função psicométrica não foi modulada pela ausência ou presença de informações de profundidade, sejam elas provenientes do

gradiente de perspectiva ou do gradiente fotografado. No entanto, os estímulos foram mais bem discriminados em tempos de exposição maiores. Este resultado está em acordo com as expectativas, pois a aquisição e processamento das informações tornam-se mais eficientes e completos à medida que há mais tempo para realizar estes processos. Desse modo, o desempenho em discriminar o tamanho será melhor aos 200 ms do que aos 50 ms, ainda mais porque neste último tempo é possível captar apenas 75% da informação presente em uma cena (Salthouse & Ellis, 1980).

Os estímulos foram apresentados pelo método psicofísico das escadas duplas e a inclinação da função psicométrica foi calculada através de dois pontos desta (50% e 70,7%). Estudos anteriores indicam que este procedimento resultou em valores de inclinação mais baixos (Bernardino, Sousa & Fukusima, 2010) e com muita variabilidade (Leek, 2001) em comparação com os valores obtidos através de outros métodos psicofísicos. No entanto, as inclinações observadas não diferiram fortemente das encontradas por Bernardino (2008), que utilizou o método dos estímulos constantes e uma tarefa semelhante à deste estudo.

No que diz respeito às distorções de tamanho, observou-se a existência de diferenças temporais no processamento das informações de tamanho e profundidade em função da maneira como o espaço foi representado. Watanabe (2006) havia indicado que as representações do espaço através de desenhos gerados computacionalmente ou de fotografias são processadas de modo semelhante, porém nada mencionou sobre o tempo de processamento destas. Desse modo, nossos resultados são relevantes e fornecem uma importante contribuição às investigações nesta área, pois a dinâmica temporal destes processos ainda era desconhecida.

Na condição com gradiente de perspectiva, as maiores distorções de tamanho foram observadas aos 100 ms, revelando que as informações de tamanho e profundidade são processadas e integradas em tempos não superiores a este. Resultados semelhantes foram descritos por Champion e Warren (2008). Com uma tarefa de busca visual, os autores investigaram a hipótese de que o tamanho retiniano é acessado diretamente em tarefas que requerem um processamento visual muito rápido. Os resultados indicaram que a percepção de tamanho relativo ocorreu em torno dos 100 ms, achado este que os levou a refutar sua hipótese e a concluir que não temos acesso explícito à representação do tamanho na retina. Tal conclusão também encontra respaldo no estudo de imageamento cerebral conduzido por

Murray et al. (2006), que mostrou que a atividade em V1 está associada ao tamanho percebido e não ao tamanho retiniano.

Por outro lado, o gradiente fotografado modulou de maneira mais intensa a percepção de tamanho em um tempo de exposição maior (150 ms). As distorções de tamanho observadas nesta condição são semelhantes às encontradas no gradiente de perspectiva aos 100 ms. Este resultado aponta que o gradiente de textura natural em questão requereu mais tempo para ser eficientemente processado e, assim, afetar o tamanho dos estímulos em comparação com o gradiente de perspectiva. Uma hipótese para explicar este resultado é a diferença na complexidade destes gradientes, que pode ser indicada pela quantidade de informações de profundidade contida neles.

O gradiente fotografado era composto de pequenas pedras arranjadas em fileiras horizontais e verticais. Pode-se afirmar, portanto, que este gradiente encerrava as três informações de textura descritas por Hillis et al (2004): perspectiva, compressão e densidade. Assim, haveria a necessidade do envolvimento de mais recursos cerebrais no processamento do gradiente natural fotografado, em que há informações visuais mais complexas. Por isso, a representação do espaço nesse caso estaria mais sujeita à influência de processos cognitivos e mecanismos *top-down*. Em contraste, o gradiente de perspectiva possui apenas uma destas informações e, por isso, seria menos complexo e necessitaria de um tempo menor para ser processado.

O resultado de Torro-Alves et al. (2012) pode ser tomado como argumento contrário à hipótese acima descrita. Estes autores mostraram que é a integração de informações de perspectiva e compressão e sua consequente influência sobre a representação de tamanho ocorreu em tempos inferiores a 100 ms. Contudo, estes resultados não invalidam a hipótese supracitada, pois eles não podem ser extrapolados e nem comparados diretamente com o nosso estudo. O primeiro motivo é que Torro-Alves et al. (2012) utilizaram apenas gradientes de textura gerados computacionalmente. Assim, a dinâmica temporal da integração de informações de perspectiva e compressão pode ser diferente em um gradiente natural fotografado. Em segundo lugar, os autores apresentaram fundos de tela que continham a linha do horizonte, além dos gradientes de textura. Sabe-se que a linha do horizonte é uma dica de profundidade e afeta fortemente a percepção do tamanho relativo (ver Ozkan & Braunstein, 2010), o que talvez contribua para a rapidez observada na integração destas informações.

A grande capacidade de nosso sistema visual em extrair informações de profundidade a partir do gradiente de textura, especialmente da perspectiva, já havia sido descrito em estudos com métodos psicofísicos (Allison & Howard, 2000; Goodenough & Gilliam, 1997; Wu, He & Ooi, 2007) e abordagem neurofisiológica (Di Lollo et al., 2000; Tong 2003). Nossos resultados extrapolam os achados anteriores ao mostrar que a rapidez e eficiência para processar os gradientes de textura se refletem também na modulação do tamanho percebido.

Uma pergunta emerge: os resultados encontrados neste experimento dão suporte a HITD ou ao modelo da percepção direta de Gibson? Não é possível responder a esta questão, pois não foram realizadas manipulações diretas para verificar a adequação destas teorias. Em um estudo recente, Tozawa (2010) realizou-as e encontrou que o modelo de percepção direta foi mais adequado que a HITD para prever o tamanho relativo. Pode-concluir, contudo, que a rapidez com que as informações de tamanho e profundidade são processadas e integradas indica a primazia dos mecanismos *bottom-up*.

Cabe ressaltar, entretanto, que a simples associação entre processos *bottom-up* e a teoria de Gibson é controversa. Higashiyama e Shimono (2004) defendem que a informação de profundidade é registrada antes que a de tamanho, o que seria consonante com a HITD, porém também afirmam que este é um processo *bottom-up*. Além disso, Norman (1980) indica que quanto mais difícil a tarefa, maior é o envolvimento de mecanismos mediativos ou *top-down*. No experimento aqui descrito, a dificuldade da tarefa de discriminação é manifesta pela pequena diferença entre o tamanho dos estímulos: a menor diferença entre os estímulos era 1 mm (ou 0,1° de ângulo visual) e a maior era 5 mm (0,5°). Portanto, os resultados também podem ser tomados como evidências em favor à HITD. Torna-se evidente, portanto, a complexidade deste tema e a dificuldade de chegar-se a uma conclusão definitiva.

E o que ocorreu nas comparações de tamanho realizadas na condição controle, isto é, sem gradiente de textura? Em todos os tempos de exposição investigados, observou-se uma superestimação entre 1 e 2% aproximadamente. Esta pequena distorção no tamanho percebido é creditada à influência da altura no campo visual, que é uma informação de profundidade também (Epstein, 1966). Entretanto, seu efeito é mais sutil do que o de um gradiente de textura. Por exemplo, Tozawa (2010) encontrou que o gradiente de textura afeta a percepção de tamanho duas vezes mais do que a altura no campo visual. Provavelmente esta diferença acentua-se nos estágios iniciais do processamento, o que explicaria os resultados do

Experimento 1 que apontam distorções de tamanho até 4 vezes maiores nas duas condições com gradiente de textura em comparação com a condição controle.

Ademais, observou-se que as distorções de tamanho aos 50 ms são similares tanto na ausência quanto na presença dos gradientes de textura, sejam estes gerados no computador ou fotografados. Este resultado indica que as informações de profundidade não modularam diferencialmente o tamanho percebido, sugerindo que estas não foram processadas de maneira eficiente. De acordo com Tsutsui et al. (2002), o gradiente de textura não é um indício de profundidade *per se*, o que nos leva a concluir que 50 ms não é um tempo suficiente para processar ou mesmo detectar os gradientes de textura em questão. Neste ponto, não se pode negligenciar um problema acerca da precisão das apresentações dos estímulos por 50 ms. O monitor utilizado possuía uma frequência de atualização de 85 Hz, o que gerava 4 ou 5 telas e, por isso, exposições de 48 ou 60 ms, respectivamente. No entanto, tal evidência não invalida as conclusões expostas, apenas situa-as em uma faixa temporal e não em um valor exato.

Os resultados aqui apresentados também mostraram uma diminuição nos desvios das respostas depois dos 100 ms para o gradiente de perspectiva e após os 150 ms para o gradiente fotografado. Uma hipótese para explicar estes dados é a aplicação de alguma correção cognitiva à medida que o tempo aumenta, por exemplo, a alocação da atenção. Assim, a atenção melhoraria o desempenho dos participantes nesta tarefa, resultando em distorções de tamanho menores à medida que se aumenta o tempo de exposição dos estímulos. No entanto, ainda falta estabelecer a exata influência da atenção sobre a percepção de tamanho, pois alguns trabalhos concluem que a atenção reduz o tamanho percebido (Tsal et al., 2005), enquanto outros afirmam o contrário (Masin 2003).

Os movimentos oculares também podem estar envolvidos na redução das distorções de tamanho. Por exemplo, Richards e Miller (1971) encontraram que a ocorrência de movimentos oculares reduziu em 18% a magnitude de uma versão da ilusão de Ponzo⁵. Além disso, não se pode negligenciar que estes movimentos ocorrem amiúde em tempos iguais ou superiores a 100 ms (Bourne, 2006) e também compartilham vias neurais com a atenção.

Em função dessas últimas observações, planejou-se o Experimento 2 com o objetivo de verificar a ocorrência de movimentos oculares nos tempos de 150 ms e 200 ms em uma tarefa de comparação de tamanho sob gradientes de textura. Esperava-se, assim, determinar se mecanismos cognitivos são responsáveis pela diminuição nos desvios de resposta à medida que aumenta o tempo de exposição dos estímulos, como foi observado no Experimento 1.

⁵ Ilusão visual geométrica, na qual duas linhas paralelas de tamanhos iguais são percebidas como de tamanhos diferentes ao estarem circundadas por duas linhas convergentes. A linha mais próxima do ponto de convergência é percebida como maior.

Além disso, realizou-se uma condição com um tempo maior de exposição (1500 ms) para identificar os mecanismos *bottom-up* e *top-down* envolvidos no processamento das informações de tamanho e profundidade.

4. Experimento 2

4.1 Objetivo

Investigar a ocorrência de movimentos oculares em tempos breves de exposição (150 ms e 200 ms) em uma tarefa semelhante à do Experimento 1. Ademais, realizou-se uma condição com menor restrição temporal (1500 ms) para identificar o padrão dos movimentos oculares neste contexto.

4.2 Método

4.2.1 Participantes

Participaram deste experimento 24 voluntários (12M e 12F) distribuídos aleatoriamente em 3 grupos equitativos, de acordo com a condição experimental a qual foram designados. Os participantes eram alunos da Universidade de Barcelona, Espanha, com idade entre 18 e 32 anos, de ambos os sexos, com acuidade visual normal ou corrigida em ambos os olhos. Este experimento foi aprovado pelo comitê de ética institucional da Universidade de Barcelona, de acordo com a Declaração de Helsinki de 1964 e suas versões posteriores (ANEXO D).

4.2.2 Material e equipamento

As condições de exposição breve dos estímulos (150 e 200 ms) foram realizadas em um sistema de duas telas (*dual screen*) composto por dois computadores: 1) Toshiba (15,4”), Intel Core 2 Duo Processor T6400, 4GB SDRAM, plataforma Windows XP, sob resolução de 1024 x 768 pixels e frequência de atualização de 60 Hz, com o programa Superlab 4® para executar o experimento e coletar as respostas; e 2) Dell (22”), Intel i7 930, 8GB RAM, plataforma XP, sob resolução de 1024 x 768 pixels e frequência de atualização de 60 Hz, com o programa SMI Experiment Suite 360°™ para apresentar os estímulos (ver Figura 11). Na condição com tempo de exposição de 1500 ms, utilizou-se apenas o segundo computador e

seu respectivo programa para apresentar os estímulos e coletar as respostas. As respostas foram dadas através de um teclado numérico à parte.

O registro dos movimentos oculares foi realizado através do SMI (SensoMotoric Instruments GmbH®) iView x RED500, com uma taxa de registro de 500 Hz. O equipamento tem uma acurácia de $0,4^\circ$ de ângulo visual e funciona em um *dark pupil system*. Neste sistema o olho é iluminado por uma luz infravermelha e há uma câmera sensível a esta luz registrando os olhos do participante. O olho e a face refletem a maior parte desta iluminação, porém a pupila a absorve, gerando uma imagem como uma elipse escura de alto contraste. Um programa computacional de análise de imagem determina o centro da pupila e a posição é mapeada através de um algoritmo. Este equipamento também compensa pequenos movimentos de cabeça (50 cm/s) ao monitorar o reflexo da córnea, cuja posição muda com movimentos da cabeça em relação à câmera, e utilizar sua posição juntamente com a localização da pupila para determinar a direção do olhar.

Os participantes posicionavam a cabeça a 60 cm do monitor e esta distância era mantida constante por um apoiador de queixo. Os estímulos e os fundos de tela utilizados foram os mesmos e encerravam ângulos visuais iguais aos do Experimento 1.

Figura 11 - Foto do equipamento e do sistema de duas telas (*dual screen*) utilizado para a execução do experimento nas condições breves de apresentação dos estímulos (150 e 200 ms). O *eye tracker* SMI iView x RED500 está localizado abaixo do monitor de 22 polegadas.

4.2.3 Procedimento

Em sessão individual, os participantes sentaram em frente ao monitor a uma distância de 60 cm com a cabeça no apoiador de queixo. O pesquisador fornecia as instruções da tarefa em uma folha de papel impressa (ANEXO E) e, após o participante terminar a leitura, certificava-se de não haver dúvidas na execução desta. A sessão experimental iniciava com um treino para a tarefa com 10 tentativas. Depois era realizada a calibração do *eye tracker*, etapa fundamental para garantir a precisão dos registros. Esta era feita com a apresentação de um ponto móvel que assumia 9 posições diferentes na tela do computador e o participante deveria seguir este ponto com os olhos. Se a calibração indicasse uma acurácia espacial pior que 1° de ângulo visual em qualquer um dos 9 pontos, esta era repetida quantas vezes fossem necessárias para atingir a precisão supracitada.

Após esta etapa, havia a validação para certificar que a calibração foi adequada. Novamente um ponto movia-se pelo monitor ocupando 5 localizações distintas e, caso a acurácia espacial fosse pior que 1° de ângulo visual em algum dos 5 pontos, a validação era repetida até esta precisão ser alcançada.

Para cada observador, os 11 pares de estímulos foram apresentados aleatoriamente uma única vez em 3 sessões diferentes, uma para cada um dos tempos de exposição (150, 200 e 1500 ms), e sempre sobre um mesmo fundo de tela (controle, gradiente de perspectiva e gradiente fotografado). A tarefa dos participantes foi a mesma do Experimento 1, isto é, indicar qual círculo apresentado era o maior.

Em cada tentativa era apresentado um ponto de fixação por 1500 ms, seguido da tela com os estímulos e o gradiente de textura por 150, 200 ou 1500 ms, dependendo da sessão experimental. Depois era apresentada uma tela para o participante emitir sua resposta, esta tela tinha uma duração de 1000 ms. Nos tempos de 150 e 200 ms, após a resposta havia uma tela em branco para que o participante controlasse a apresentação do próximo estímulo mediante o pressionar de uma tecla. Para o tempo de 1500 ms, não havia esta tela, assim após a resposta, dava-se início à tentativa seguinte. As sessões foram realizadas no mesmo dia e o intervalo entre elas era definido pelo próprio participante. O tempo total médio do experimento foi de 20 minutos.

4.2.4 Análise de dados – Respostas

A porcentagem de erros cometidos nas comparações de tamanho em cada fundo de tela (controle, gradiente de perspectiva e gradiente fotografado) e em cada tempo de exposição (150, 200 e 1500 ms) foi utilizada como índice do desempenho dos participantes na tarefa. Foi considerado erro quando os participantes indicavam um estímulo de tamanho menor como sendo o maior naquela tentativa. Assim, para este cômputo foram excluídas as tentativas em que o estímulo teste e o estímulo padrão tinham o mesmo tamanho (100 *pixels*).

Os erros de julgamento foram submetidos aos testes não-paramétricos de Kruskal-Wallis com um fator independente (“tempo de exposição”) e de Friedman com um fator de medida repetidas (“fundo de tela”). Foram realizados testes qui-quadrado para analisar a distribuição dos erros, de acordo com o tamanho do estímulo teste, e também testes binomiais para verificar se o desempenho dos participantes era superior ao de respostas ao acaso.

Os tempos de resposta para comparações realizadas aos 1500 ms foram submetidos a uma ANOVA com um fator independente (“fundo de tela”). A ANOVA foi realizada somente após as exigências de normalidade e de homocedasticidade das variâncias serem verificadas através dos testes de Shapiro-Wilk e de Levene, respectivamente. As possíveis diferenças encontradas no efeito principal foram verificadas através do pós-teste de Bonferroni. A análise estatística foi realizada com o auxílio do programa SPSS 17.0.

4.2.5 Análise de dados – Parâmetros oculares

As fixações foram identificadas pelo critério de Salvucci e Goldberg (2000), ou seja, era necessário que o olhar do participante se mantivesse por pelo menos 100 ms em uma área menor que 1° de ângulo visual. As sacadas foram definidas por uma mudança na posição espacial de pelo menos 0,5°, uma velocidade mínima de 35°/s e uma aceleração mínima de 9500°/s². Os registros do *eye tracker* SMI iView RED500TM foram pré-analisados com o programa SMI BeGazeTM.

Duas estratégias foram empregadas para permitir a comparação estatística entre estes parâmetros. A primeira foi a definição de áreas ou regiões de interesse (*area of interest*, AOI na sigla em inglês) nas telas apresentadas aos participantes. Estabeleceu-se 4 AOIs, a saber: 1) uma na região central, coincidindo com a posição do ponto de fixação (PF); 2) uma para o círculo acima (CAc); 3) uma para o círculo abaixo (CAb); e 4) uma para o restante da imagem, a qual foi denominada fundo (F). Um esquema destas AOIs pode ser visto na Figura

12. A segunda estratégia utilizada foi verificar se os parâmetros investigados variavam em função do tamanho dos círculos (maior e menor), independente da localização espacial destes.

Figura 12 - Um esquema das áreas de interesse (AOIs) estabelecidas para a análise dos parâmetros oculares. A área 1 coincide com a localização do ponto de fixação (PF), no qual o participante deve manter o olhar até a apresentação dos estímulos. O número 2 indica a área do círculo apresentado acima (CAc), o 3 a área do círculo apresentado abaixo (CAb) e a área 4 é composta pelo restante da imagem, consistindo no fundo de tela (F).

A análise dos parâmetros oculares foi realizada separadamente em função do tempo de exposição. Decidiu-se por essa divisão, dada a diferença dos objetivos nas condições com tempos curtos (150 e 200 ms) e longo (1500 ms) de apresentação dos estímulos. Nos tempos breves de exposição verificou-se a porcentagem de tentativas em que ocorreram movimentos oculares, a localização espacial destes e se guardavam relação com o desempenho dos participantes na tarefa.

No tempo de 1500 ms, teve-se por objetivo identificar componentes perceptuais e cognitivos envolvidos no processamento dos estímulos para a realização da tarefa de comparar tamanhos. Para isso, as seguintes medidas espaciais e temporais das fixações e dos movimentos sacádicos foram analisadas:

- o número total e médio das fixações, a localização espacial destas e sua sequência temporal através do teste qui-quadrado;
- a duração média das fixações e a duração média da primeira fixação através de uma ANOVA com um fator independente (“fundo de tela”);

- duração total média das fixações dirigidas a uma mesma AOI (*dwell time*) através de uma ANOVA mista com um fator independente (“fundo de tela”) e um fator de medida repetida (“AOI”). O teste t pareado foi utilizado para se analisar os fundos de tela separadamente;

- a direção do primeiro movimento sacádico através do teste qui-quadrado;

- a amplitude média dos movimentos sacádicos através de uma ANOVA com um fator independente (“fundo de tela”). O teste t pareado foi utilizado para se analisar os fundos de tela separadamente.

As ANOVAs supracitadas foram realizadas somente após as exigências de normalidade e de homocedasticidade das variâncias serem verificadas através dos testes de Shapiro-Wilk e de Levene, respectivamente. As possíveis diferenças encontradas nos efeitos principais foram verificadas através do pós-teste de Bonferroni. A análise estatística foi realizada com o auxílio do programa SPSS 17.0.

4.3 Resultados

4.3.1 Erros de julgamento

As porcentagens dos erros cometidos nas comparações de tamanho foram submetidas a dois testes estatísticos não-paramétricos. Primeiramente, para verificar se havia diferença nos erros de julgamento em cada tempo de exposição, em função do fundo de tela, empreendeu-se o teste de Kruskal-Wallis. A quantidade de erros cometidos nos tempos de 150 ms ($\chi^2 = 0,48$, $p = 0,785$) e de 1500 ms ($\chi^2 = 0,01$, $p = 0,997$) foi semelhante em todos os fundos de tela. No entanto, os gradientes de textura afetaram diferencialmente os erros de julgamento cometidos nas apresentações de 200 ms ($\chi^2 = 11,80$, $p = 0,003$). Nota-se pela Figura 13 que os participantes cometeram mais erros na condição controle ($n = 30\%$) do que no gradiente de perspectiva ($n = 15\%$) e no gradiente fotografado ($n = 12,5\%$).

Figura 13 - Porcentagem dos erros de julgamentos nas comparações de tamanho realizadas nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado) e nos 3 tempos de exposição (150, 200 e 1500 ms). As barras de erro referem-se ao erro padrão da média. Os participantes observavam os estímulos em apenas um dos fundos de tela (medida independente) e nos três tempos de exposição (medida repetida).

Em seguida, utilizou-se o teste de Friedman para verificar se as porcentagens dos erros cometidos em cada fundo de tela diferiam em função dos tempos de exposição dos estímulos. Os resultados indicaram que o tempo de exposição foi fator importante na quantidade de erros de julgamento cometidos, independente do fundo de tela, a saber: controle ($\chi^2 = 15,72$, $p < 0,001$), gradiente de perspectiva ($\chi^2 = 10,46$, $p = 0,005$) e gradiente fotografado ($\chi^2 = 12,56$, $p = 0,002$). Fica evidente pela observação da Figura 13, que menos erros foram cometidos no tempo de exposição de 1500 ms, em comparação com as exposições breves (150 e 200 ms). Este resultado mostra que um maior tempo de exposição aumenta o envolvimento de recursos de processamento de ordem superior, o que se refletiu em melhor desempenho na tarefa.

Na Tabela 1 é apresentada a quantidade de erros de julgamento cometidos em cada um dos estímulos teste, em função do tempo de exposição. Cabe ressaltar que os dados dos 3 fundos de tela foram reunidos. Tomando-se os tempos de exposição em conjunto, o teste qui-quadrado revelou que a distribuição dos erros não foi equitativa ($\chi^2 = 27,72$, $p = 0,001$). Observaram-se mais erros nas tentativas em que o tamanho do estímulo teste era mais próximo (97 e 103 *pixels*) do tamanho do estímulo padrão (100 *pixels*), o que pode ser conferido na coluna mais à direita da Tabela 1.

Ao analisarem-se os tempos de exposição separadamente, observou-se que a quantidade de erros foi igualmente distribuída pelos tamanhos do estímulo teste nos tempos de 150 ($\chi^2 = 11,63$, $p = 0,235$) e 200 ms ($\chi^2 = 9,22$, $p = 0,417$). Por outro lado, no tempo de 1500 ms foi revelada uma distribuição desigual nos erros ($\chi^2 = 15,80$, $p = 0,417$), sendo que os participantes cometiam mais erros de julgamento no estímulo teste de 97 *pixels* ($n = 8$).

Tabela 1 - O número de erros de julgamentos dos participantes nos 3 tempos de exposição (150, 200 e 1500 ms), em função dos 11 tamanhos do estímulo teste.

Tamanho do estímulo teste (<i>pixels</i>)	Tempo de exposição (ms)			
	150	200	1500	TOTAL
85	3	3	0	6
88	7	2	0	9
91	3	5	1	9
94	3	6	1	10
97	9	9	8	26
103	11	7	3	21
106	9	3	1	13
109	5	4	1	10
112	8	4	0	12
115	6	3	0	9
TOTAL	64	46	15	125

Por fim, analisou-se também se o desempenho dos participantes na tarefa foi diferente de respostas ao acaso. Isto é, os participantes discriminaram as diferenças de tamanho entre os estímulos ou indicaram sua resposta aleatoriamente? Conforme revelado por testes binomiais ($p < 0,001$), as respostas são diferentes do que seria esperado caso os observadores respondessem ao acaso. Este resultado indica que independente do gradiente de textura utilizado, nosso sistema visual é capaz de extrair informações corretas de tamanho mesmo em tempos curtos de apresentação dos estímulos.

4.3.2 Tempos de exposição de 150 e 200 ms

Em cada tentativa realizada pelos participantes, verificou-se se havia alguma fixação fora da área PF. Este critério foi escolhido porque a instrução indicava que o participante mantivesse o olhar no ponto de fixação até o aparecimento dos estímulos. Em seguida, ele

poderia movimentar os olhos livremente. Desse modo, esperava-se que as fixações estivessem localizadas na área PF, onde o participante estava olhando anteriormente.

A porcentagem de tentativas em que ocorreram movimentos oculares é apresentada na Tabela 2, de acordo com o tempo de exposição e o fundo de tela. Nota-se que em todas as condições, os movimentos oculares ocorreram em menos de 10% das tentativas. Tal resultado indica que mecanismos cognitivos não explicam totalmente a diminuição dos desvios de resposta a partir dos 150 ms, observada no Experimento 1. Destes dados ainda depreende-se que a ocorrência de movimentos oculares foi menor no gradiente fotografado em comparação com os outros dois, inclusive não sendo registrado nenhuma sacada no tempo de 150 ms.

Tabela 2 - A porcentagem (%) de tentativas em que ocorreram movimentos oculares, em função dos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado) e dos 2 tempos de exposição (150 e 200 ms).

	Controle	Gradiente de perspectiva	Gradiente fotografado
150 ms	4,35	7,61	0
200 ms	9,78	8,70	5,43

Neste momento, uma pergunta emerge: os movimentos oculares, ainda que poucos, seguiram algum padrão? Para comparações de tamanho realizadas aos 150 ms foram observadas apenas 11 movimentos sacádicos somando-se os 3 fundos de tela investigados. Não se observou uma diferença significativa pelo teste qui-quadrado em relação à localização espacial final destes (5 para CAc, 3 para CAb e 2 para F; $\chi^2 = 1,40$, $p = 0,497$) e tampouco em função do tamanho do estímulo para onde o movimento sacádico foi dirigido (3 para o maior círculo e 5 para o menor; $\chi^2 = 0,50$, $p = 0,480$).

Os resultados foram diferentes para comparações de tamanho realizadas aos 200 ms. Dos 22 movimentos sacádicos observados, 17 foram dirigidos ao CAc. Pelo teste qui-quadrado ($\chi^2 = 19,18$, $p < 0,001$), este número é significativamente maior do que os direcionados ao CAb ($n = 3$) e ao F ($n = 2$). Novamente, nenhum padrão pode ser extraído em função do tamanho do estímulo, que foi alvo do movimento sacádico (10 para o maior círculo e 8 para o menor; $\chi^2 = 0,22$, $p = 0,637$).

Além disso, analisou-se se havia alguma relação evidente entre os movimentos oculares e o desempenho na tarefa. Foram observados mais acertos do que erros nas tentativas em que ocorreram movimentos oculares (80% de acertos aos 150 ms e 65% aos 200 ms) e o

teste binomial indicou ser significativa tal diferença ($p < 0,001$ e $p = 0,004$, respectivamente). Assim, pode-se afirmar que os movimentos oculares predisseram o acerto nas tentativas em que foram observados.

4.3.3 Tempo de exposição de 1500 ms

4.3.3.1 Tempo de resposta

A análise do tempo de resposta dos participantes não estava prevista inicialmente, porém o pesquisador observou durante a sessão experimental, que os participantes emitiam suas respostas antes que a apresentação dos estímulos fosse interrompida aos 1500 ms. É interessante notar, que nenhuma instrução para que respondessem o mais rápido possível foi dada. Assim, como o tempo de resposta reflete o processamento perceptivo inicial, esta análise poderia fornecer informações adicionais sobre a rapidez com que as informações de tamanho e profundidade são processadas.

Os dados do tempo de resposta foram submetidos a uma ANOVA de um fator “fundo de tela” como variável independente. A ANOVA revelou um efeito marginal dos fundos de tela sobre o tempo de resposta ($F_{2,23} = 2,855$, $p = 0,080$). Observou-se uma tendência de maior facilidade para o processamento do gradiente fotografado ($\bar{x} = 1103,78$, $EPM = 142,57$) em comparação com a condição controle ($\bar{x} = 1821,40$, $EPM = 294,87$) e com o gradiente de perspectiva ($\bar{x} = 1587,02$, $EPM = 182,82$).

4.3.3.2 Número de fixações

Na Tabela 3 é apresentado o número total de fixações realizadas pelos participantes para cada tamanho do estímulo teste, em função do fundo de tela. Nota-se que em média foram realizadas de 4 a 5 fixações, independentemente do tamanho do estímulo teste e do gradiente de textura. Este resultado é consonante com a média de 3 a 4 fixações por segundo observada em nossa interação cotidiana com o ambiente, que foi descrita por Tatler et al. (2005).

Analizou-se o número de fixações nos 3 fundos de tela em conjunto e o teste qui-quadrado revelou que não houve mais fixações em nenhum dos 11 estímulos teste ($\chi^2 = 5,636$, $p = 0,845$). Resultados semelhantes foram observados na análise dos fundos de tela individualmente, isto é, o tamanho do estímulo teste também não influenciou o número de

fixações realizadas pelos participantes em cada gradiente de textura, a saber: sem textura ($\chi^2 = 11,68$, $p = 0,307$), gradiente de perspectiva ($\chi^2 = 4,27$, $p = 0,934$) e gradiente fotografado ($\chi^2 = 2,02$, $p = 0,996$).

Tabela 3 - O número total de fixações dos participantes para os 11 tamanhos do estímulo teste, em função dos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado) .

Tamanho do estímulo teste (<i>pixels</i>)	Fundos de tela			Total
	Controle	Gradiente de perspectiva	Gradiente fotografado	
85	38	36	35	109
88	33	42	39	114
91	34	34	43	111
94	33	37	39	109
97	37	38	41	116
100	32	39	35	106
103	39	36	41	116
106	46	33	41	120
109	32	36	36	104
112	20	31	40	91
115	33	45	36	114
Total	377	407	426	1210

Além disso, analisou-se o número total de fixações, para verificar se os gradientes de textura estudados diferiam no que concerne à quantidade de informação relativa. Assim, o número total de fixações nos respectivos fundos de tela foi submetido ao teste qui-quadrado. Este não revelou diferença significativa ($\chi^2 = 3,03$, $p = 0,220$), embora tenham sido observadas mais fixações no gradiente fotografado ($n = 426$) em comparação com o gradiente de perspectiva ($n = 407$) e deste em relação à condição controle ($n = 377$). Este resultado fornece indícios de que a quantidade de informação contida nos gradientes de textura é diferente e, provavelmente, com maior tempo de exposição esta diferença se revelaria significativa.

4.3.3.3 Dimensão espacial das fixações

É fundamental identificar as localizações espaciais das fixações, pois estas nos informam sobre as áreas onde se concentra a relevância informativa da imagem. Um teste qui-quadrado revelou que as frequências das fixações nas quatro AOIs, tomando-se os três fundos em conjunto, não foram distribuídas igualmente ($\chi^2 = 501,48$, $p < 0,001$). Observaram-se poucas fixações localizadas na área F ($n = 49$), que corresponde ao fundo de tela (gradiente de textura ou ausência deste); e uma maior concentração das fixações na área PF, região central ($n = 586$). Primeiramente, estes resultados mostram que os participantes não necessitaram mover os olhos para o fundo de tela para executar a tarefa. Em seguida, pode-se afirmar que houve uma tendência dos participantes em dirigir e/ou manter os olhos na região do ponto de fixação, o que está em acordo com os estudos de Tatler et al. (2005) e Tatler (2007).

Optou-se então por excluir as áreas PF e F e verificar se houve diferença entre as frequências das fixações, de acordo com a localização espacial dos círculos (áreas CAc e CAb). O teste qui-quadrado mais uma vez revelou que as fixações não foram igualmente distribuídas ($\chi^2 = 23,00$, $p < 0,001$), sendo que o círculo apresentado no campo visual superior recebeu mais fixações ($n = 345$) do que o círculo apresentado no campo visual inferior ($n = 230$). Outra análise foi realizada para identificar se houve maior número de fixações em função do tamanho dos círculos e não entre as AOIs. Novamente, o teste qui-quadrado revelou uma distribuição desigual das fixações ($\chi^2 = 23,00$, $p < 0,001$), com mais fixações localizadas no círculo maior ($n = 318$) em comparação com o círculo menor ($n = 208$).

Decidiu-se também empreender as mesmas análises para cada gradiente de textura separadamente. Resultados semelhantes foram encontrados no que concerne às AOIs. Em todos os fundos de tela, através do teste qui-quadrado, observou-se uma distribuição desigual das fixações ($p < 0,001$). Estas tiveram uma maior concentração na área PF e poucas localizadas na área F, independentemente do fundo de tela, como pode ser observado na Tabela 4. Também foi revelado um maior número de fixações no círculo acima quando comparado ao círculo abaixo após a exclusão das áreas PF e F ($p = 0,027$ para o fundo controle; $p = 0,013$ para o gradiente de perspectiva; e $p < 0,001$ para o fundo fotografado).

Tabela 4 - O número total de fixações dos participantes nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado), em função das 4 áreas de interesse (PF, CAC, CAb e F).

Fundo de tela	Número de fixações nas áreas de interesse (AOIs)				Total
	PF	CAC	CAb	F	
Controle	196	101	72	8	377
Gradiente de perspectiva	182	122	86	17	407
Gradiente fotografado	208	122	72	24	426
Total	586	345	230	49	1210

Por outro lado, ao considerar-se o tamanho dos círculos, a prevalência de fixações no círculo maior não foi observada em todos os fundos de tela. No fundo sem textura ($\chi^2 = 24,32$, $p < 0,001$) e no gradiente de perspectiva ($\chi^2 = 5,08$, $p < 0,024$), observou-se uma diferença significativa nas fixações localizadas no círculo maior e no círculo menor, com este recebendo menos fixações do que aquele (ver Tabela 5). No gradiente fotografado, entretanto, apesar de mais fixações estarem localizadas no círculo maior, esta diferença não é significativa ($\chi^2 = 1,62$, $p = 0,204$).

Tabela 5 - O número total de fixações dos participantes nos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado), em função do tamanho dos círculos (maior e menor).

Fundo de tela	Número de fixações nos círculos em função do tamanho		Total
	Maior	Menor	
Controle	110	48	158
Gradiente de perspectiva	110	79	189
Gradiente fotografado	98	81	179
Total	318	208	526

4.3.3.4 Sequência das fixações

Na tabela 6 está representada a frequência da localização espacial das fixações nas AOIs para cada fundo de tela, em função da sequência em que ocorreram. É importante salientar que a posição da primeira fixação não foi mostrada por sempre ser dirigida à área PF. Decidiu-se também apresentar somente até a sexta fixação, pois em média os participantes realizaram de 4 a 5 fixações em cada tentativa.

Estas frequências foram submetidas ao teste qui-quadrado. Cabe ressaltar que a área F foi excluída das análises, dada a baixa frequência com que esta área foi alvo do olhar dos participantes em qualquer posição na ordem das fixações. Em todos os fundos de tela, a segunda e a terceira fixação foram preferencialmente dirigidas à área PF ($p < 0,001$). Como a primeira fixação foi sempre localizada na área PF, não surpreende que a segunda fixação tenha sido dirigida à mesma área. Na terceira fixação, a área PF continuou recebendo mais fixações, embora tenha aumentado a quantidade de fixações na área CAC principalmente, mas também na área CAB.

De maneira geral, a partir da quarta fixação, os participantes distribuíram suas fixações entre as áreas PF, CAC e CAB de maneira mais equitativa. Apenas no gradiente de perspectiva ainda observou-se uma distribuição desigual entre as AOIs na quarta fixação ($\chi^2 = 11,74$, $p < 0,003$), inclusive com mais fixações na área CAC ($n = 28$) do que na área PF ($n = 26$) e poucas fixações dirigidas à área CAB ($n = 8$).

Além disso, cabe ressaltar que mesmo a diferença não sendo significativa ($p > 0,001$), observou-se que a quarta e a quinta fixação no gradiente de perspectiva e a quinta e a sexta fixação no gradiente fotografado foram mais dirigidas à CAC do que à área PF. Por outro lado, no fundo controle, o teste qui-quadrado revelou uma distribuição desigual das fixações entre as AOIs na sexta fixação ($\chi^2 = 8,15$, $p < 0,017$), com a área CAB recebendo mais fixações ($n = 19$) que as áreas PF ($n = 16$) e CAC ($n = 5$).

Conclui-se que os participantes utilizaram uma estratégia sequencial muito semelhante em sua conduta exploratória da imagem e que o fundo de tela foi uma informação pouco relevante para alterar tal estratégia nas fixações iniciais. No entanto, os resultados também evidenciam que a localização espacial em torno da quinta ou sexta fixação foi modulada diferencialmente pela presença ou não de informação de profundidade.

Tabela 6 - A frequência da localização espacial das fixações nas 4 AOIs (PF, CAc, CAb e F) para os 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado), em função da sequência em que ocorreram até a sexta fixação.

		Ordem das fixações				
		2 ^a	3 ^a	4 ^a	5 ^a	6 ^a
Controle	PF	69	45	24	21	16
	CAc	8	22	23	16	5
	CAb	5	14	19	15	19
	F	1	1	1	2	2
Gradiente de perspectiva	PF	56	40	26	20	17
	CAc	14	21	28	24	16
	CAb	13	14	8	11	10
	F	6	1	4	1	2
Gradiente fotografado	PF	72	47	28	21	18
	CAc	6	16	21	27	19
	CAb	5	11	17	12	18
	F	5	4	3	3	3

4.3.3.5 Dimensão temporal das fixações

A duração de uma fixação fornece indícios sobre a quantidade de informação a ser processada em uma imagem ou parte desta. Isto é, o tempo gasto inspecionando uma área de uma imagem é diretamente proporcional à complexidade desta (Henderson, 1992). Primeiramente, verificou-se se os gradientes de textura afetavam a duração média das fixações. Para isso, realizou-se uma ANOVA com o fator “fundo de tela” como variável independente, a qual não revelou diferença significativa ($F_{2,21} = 0,172$, $p = 0,843$). Este resultado indica que, de maneira geral, os participantes necessitavam da mesma quantidade de tempo para processar a informação contida em todos os fundos de tela. Analisou-se também a duração da primeira fixação. Uma ANOVA semelhante à descrita acima revelou que os fundos de tela não afetaram os valores médios da duração das primeiras fixações ($F_{2,21} = 0,500$, $p = 0,614$).

Outra medida temporal das fixações comumente analisada é a duração total de todas as fixações em uma mesma AOI (*dwel time*, em inglês). Acredita-se que este parâmetro é mais adequado para extrair conclusões acerca do processamento cognitivo subjacente, enquanto que a duração das fixações individuais informaria sobre o processamento das características perceptivas da imagem (Henderson, 1993).

Inicialmente, a duração total média das fixações foi submetida a uma ANOVA mista de modelo: 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado) x 4 AOIs (PF, CAC, CAb e F), sendo o fator “fundo de tela” de medida independente e o fator “AOI” de medida repetida. A exigência de esfericidade foi violada ($p = 0,006$) e, por essa razão, utilizou-se o épsilon de *Greenhouse-Geisser* para ajustar os graus de liberdade. Não foi observada interação entre os fatores “fundo de tela” e “AOI” ($F_{4,47} = 0,250$, $p = 0,924$) e o tempo gasto em cada AOI não diferiu de acordo com o fundo de tela sob o qual as comparações foram realizadas ($F_{2,21} = 0,308$, $p = 0,738$). Por outro lado, o tempo que o olhar dos participantes permaneceu em cada AOI foi diferente ($F_{2,47} = 12,536$, $p < 0,001$).

O pós-teste de Bonferroni revelou que a área F foi inspecionada por menos tempo ($\bar{x} = 388,95$ ms, EPM = 140,22) em comparação com as demais áreas: PF ($\bar{x} = 4152,46$ ms, EPM = 584,08), CAC ($\bar{x} = 2648,24$ ms, EPM = 463,10) e CAb ($\bar{x} = 1864,91$ ms, EPM = 394,21). Estas diferenças significativas reveladas pelo pós-teste podem ser observadas na Figura 14. A duração total média das fixações também foi significativamente diferente entre as áreas PF e CAb ($p = 0,046$), porém optou-se por não representar esta diferença na Figura 14 com o objetivo de preservar a clareza do mesmo. Pode-se afirmar que os participantes olharam por mais tempo à área PF e dedicaram menos tempo à área F, resultado este similar ao da localização espacial das fixações.

Figura 14 – Duração total média das fixações e dos movimentos sacádicos (em ms) em cada uma das quatro áreas de interesse (PF, CAC, CAb e F), em função dos três fundos de tela (controle, gradiente de perspectiva e gradiente fotografado). As barras de erro referem-se ao erro padrão da média. As diferenças significativas reveladas pelo pós-teste de Bonferroni estão representadas na Figura.

Além disso, uma ANOVA foi realizada excluindo-se as áreas PF e F. Esta análise revelou que os participantes não permaneceram mais tempo olhando acima ou abaixo no campo visual ($F_{1,21} = 1,420$, $p = 0,247$), não sendo significativo o efeito do fundo de tela e nem sua interação com as AOIs ($F_{2,21} = 0,297$, $p = 0,746$ e $F_{2,21} = 0,013$, $p = 0,987$, respectivamente). As mesmas análises foram empreendidas para cada fundo de tela em separado e os resultados foram iguais. Independente do fundo de tela, os participantes olharam por mais tempo à área PF e menos à área F e, excluindo-se estas áreas, não se observou maior tempo de inspeção da imagem no campo visual inferior ou superior.

A duração total média das fixações também foi investigada em função do tamanho do círculo observado (maior ou menor). Os dados foram submetidos a uma ANOVA mista de modelo: 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado) x 2 tamanho do círculo (maior e menor), sendo o fator “fundo de tela” de medida independente e o fator “tamanho do círculo” como medida repetida.

A ANOVA revelou que o tempo que os participantes olharam para o círculo maior e menor foi significativamente diferente ($F_{1,21} = 12,301$, $p = 0,002$). O pós-teste de Bonferroni indicou que o círculo maior foi inspecionado por mais tempo ($\bar{x} = 2643,61$ ms, EPM = 339,12) do que o círculo menor ($\bar{x} = 1486,10$ ms, EPM = 268,61). Este resultado indica que o tamanho dos estímulos modulou a estratégia visual dos participantes, independente se apresentado acima ou abaixo no campo visual.

Analísaram-se também os fundos de tela separadamente através do teste t pareado. No fundo controle, o círculo maior foi observado por mais tempo ($\bar{x} = 3023,06$, EPM = 612,28) do que o círculo menor ($\bar{x} = 936,35$, EPM = 149,74) e esta diferença foi significativa ($t_7 = 4,48$, $p = 0,003$). Por outro lado, não foi encontrada diferença significativa no gradiente de perspectiva ($t_7 = 0,89$, $p = 0,402$) e no gradiente fotografado ($t_7 = 1,60$, $p = 0,153$), embora a duração total das fixações no círculo maior tenha sido mais longa ($\bar{x} = 2630,59$ ms, EPM = 642,10 e $\bar{x} = 2277,16$ ms, EPM = 497,85, respectivamente) do que no círculo com o tamanho menor ($\bar{x} = 1961,14$ ms, EPM = 516,31 e $\bar{x} = 1560,82$ ms, EPM = 566,04, respectivamente).

4.3.3.6 Dimensão espacial dos movimentos sacádicos

Identificou-se qual a direção do primeiro movimento sacádico. Ou seja, após a fixação inicial, para onde o participante primeiramente movimenta os olhos? Há uma estratégia única

ou ela muda de acordo com o fundo de tela? Este parâmetro é importante, pois fornece informações sobre a estratégia visual utilizada na tarefa. E por que apenas o primeiro sacádico? Há evidências de que é equiprovável a direção dos sacádicos subsequentes ao primeiro. Isto é, eles são dirigidos equitativamente entre os objetos de uma cena, não importando a sua relevância para a tarefa e/ou sua saliência visual (Henderson et al., 1999).

Primeiramente, analisou-se se a frequência com que a primeira sacada era dirigida para cima ou para baixo nos 3 fundos de tela tomados em conjunto. O teste qui-quadrado revelou que há uma diferença significativa entre as sacadas dirigidas ao campo visual superior e inferior ($\chi^2 = 36,24$, $p < 0,001$), sendo que os participantes dirigiram seu olhar para cima ($n = 172$) em maior frequência do que para baixo ($n = 77$). Esta preferência em olhar para cima, também foi observada ao analisarem-se os fundos de tela separadamente: controle ($\chi^2 = 8,05$, $p = 0,005$), gradiente de perspectiva ($\chi^2 = 11,58$, $p = 0,001$) e gradiente fotografado ($\chi^2 = 17,61$, $p < 0,001$), como pode ser observado na Tabela 7.

Tabela 7 - O número total de vezes em que o primeiro movimento sacádico foi dirigido acima ou abaixo no campo visual, em função dos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado).

Fundo de tela	Direção do primeiro sacádico		Total
	Acima	Abaixo	
Controle	55	29	84
Gradiente de perspectiva	57	26	83
Gradiente fotografado	60	22	82
Total	172	77	249

Em seguida, investigou-se a direção da primeira sacada em função do tamanho dos círculos. Analisando-se os dados dos 3 fundos de tela, o teste qui-quadrado não apontou nenhuma preferência na direção do primeiro sacádico ($\chi^2 = 2,33$, $p = 0,127$), eles foram igualmente distribuídos ao maior ($n = 125$) e ao menor círculo ($n = 102$). Na análise individual dos fundos de tela, o teste qui-quadrado também não revelou diferenças significativas para o fundo controle ($\chi^2 = 0,117$, $p = 0,732$) e tampouco para o gradiente fotografado ($\chi^2 = 0,333$, $p = 0,564$), cujos valores são encontrados na Tabela 8. Por outro lado,

nas comparações de tamanho realizadas no gradiente de perspectiva, a primeira sacada foi preferencialmente em direção ao círculo maior ($\chi^2 = 5,88$, $p = 0,015$).

Tabela 8 - O número total de vezes em que o primeiro movimento sacádico foi dirigido ao maior ou ao menor círculo, em função dos 3 fundos de tela (controle, gradiente de perspectiva e gradiente fotografado).

Fundo de tela	Direção do primeiro sacádico		Total
	Maior	Menor	
Controle	37	40	77
Gradiente de perspectiva	48	27	75
Gradiente fotografado	40	35	75
Total	125	102	227

4.3.3.7 Amplitude dos movimentos sacádicos

Outro parâmetro investigado foi a amplitude sacádica (em graus de ângulo visual), que nos informa a extensão do campo visual que está sendo explorada. A amplitude média dos movimentos sacádicos foi submetida a uma ANOVA com o fator “fundo de tela” como variável independente, a qual não revelou diferença nas amplitudes sacádicas em função do gradiente sob o qual as comparações de tamanho foram realizadas ($F_{2,279} = 1,54$, $p = 0,217$).

As amplitudes dos movimentos sacádicos também foram analisadas em função da posição no campo visual (acima e abaixo) e do tamanho dos estímulos (círculo maior e círculo menor) para onde se dirigiam em cada fundo de tela. Os dados foram submetidos ao teste t pareado. Não houve diferença entre as amplitudes dos movimentos sacádicos dirigidos ao círculo maior e ao círculo menor em nenhum dos fundos de tela, a saber: controle ($t_{41} = 0,160$, $p = 0,874$), gradiente de perspectiva ($t_{31} = -0,286$, $p = 0,777$) e gradiente fotografado ($t_{40} = 1,380$, $p = 0,175$).

Da mesma maneira, as amplitudes sacádicas também foram semelhantes independentemente se dirigidas acima ou abaixo no fundo controle ($t_{31} = 0,299$, $p = 0,767$) e no gradiente fotografado ($t_{22} = -1,360$, $p = 0,188$). No entanto, no gradiente de perspectiva foi revelada uma diferença significativa ($t_{28} = 2,044$, $p = 0,050$). Os movimentos sacádicos dirigidos acima tiveram maior amplitude ($\bar{x} = 3,64$, $EPM = 0,21$) do que os direcionados abaixo ($\bar{x} = 2,84$, $EPM = 0,36$). Os resultados são semelhantes aos de Jansen et al. (2009), que

mostraram pouca variação na amplitude sacádica vertical durante a observação de imagens em 2D e 3D.

Além da amplitude, a duração média dos movimentos sacádicos (em ms), um parâmetro da dinâmica temporal destes, também foi submetida à semelhante análise estatística. Optou-se por não descrever estas análises pormenorizadas, pois os resultados foram idênticos, dado que estes dois parâmetros estão intimamente relacionados. Ou seja, quanto maior a amplitude do movimento sacádico, maior será a sua duração.

4.4 Discussão

O segundo experimento teve por objetivo investigar o envolvimento dos mecanismos *bottom-up* e *top-down* em uma tarefa psicofísica semelhante à do primeiro experimento, isto é, comparar o tamanho de estímulos inseridos em gradiente de perspectiva e gradiente fotografado. Primeiramente, verificou-se com que frequência ocorrem movimentos oculares na realização desta tarefa em apresentações de 150 e 200 ms. Uma alta frequência destes movimentos indicaria que mecanismos *top-down* estão implicados na diminuição da distorção de tamanho observada nestes tempos de exposição no Experimento 1. Em segundo lugar, investigou-se a dimensão temporal e espacial de fixações e movimentos sacádicos na mesma tarefa, porém com menor restrição temporal (1500 ms). Esperava-se, assim, ampliar o entendimento sobre o processamento inicial e mais tardio das informações de tamanho e profundidade.

Nas condições com tempos de exposição breves (150 e 200 ms), encontrou-se que os participantes cometiam menos erros nas tentativas em que moviam os olhos. Esta evidência sugere que a diminuição nos desvios de resposta observada no Experimento 1 ocorreu devido aos movimentos oculares. Entretanto, os resultados também indicaram poucos movimentos oculares durante a realização da tarefa. Independentemente do fundo de tela, o olhar dos participantes não permaneceu na região do ponto de fixação em menos de 10% das tentativas, evidência esta que dá suporte à idéia do *scene gist* (Biederman; Rabinowitz, Glass, Stacy, 1974; Intraub, 1980). Isto significa que nosso sistema visual extrai características globais e forma uma representação geral da imagem muito rapidamente e com uma única fixação. Além

disso, a observação de que ocorreram menos movimentos oculares nas comparações realizadas no gradiente fotografado é consonante com esta hipótese. Provavelmente, os participantes moveram menos os olhos neste fundo de tela dada a sua maior complexidade, facilitando a extração das informações globais da imagem. Estes resultados estão em acordo com a abordagem ecológica de Gibson e indicam a participação prevalente de mecanismos *bottom-up* no processamento inicial e integração das informações de tamanho e profundidade.

Na condição com menor restrição temporal (1500 ms), a despeito da análise do tempo de resposta não revelar diferença significativa, observou-se uma tendência dos participantes em responder mais rapidamente nas condições com informações de profundidade, principalmente o gradiente fotografado. Ademais, é importante ressaltar que o desempenho na tarefa foi semelhante nos 3 fundos de tela. Tomados em conjunto, estes resultados sugerem que os gradientes mobilizaram recursos atentos automáticos (*bottom-up*). Segundo Prinzmetal, Leonhardt e Garret (2008), a atenção automática não repercute sobre a acurácia, porém modifica o estágio de seleção da resposta, exatamente o que foi observado pela adição de profundidade no segundo experimento.

Com relação à dimensão temporal das fixações, observou-se que a duração média das fixações foi semelhante em todos os fundos de tela. Os resultados de Gareze e Findlay (2007) mostraram fixações mais longas em imagens com linhas desenhadas em comparação com fotografias em escala de cinza. Nosso estudo não encontrou fixações de maior duração no gradiente de perspectiva do que no gradiente fotografado, como seria esperado. No entanto, resultado semelhante ao nosso foi descrito por Jansen et al. (2009), os quais encontraram que a presença de informações de profundidade não afeta a duração das fixações, apesar destas ficarem mais breves quanto mais informações contém uma imagem. Este achado de que os gradientes de textura não modificam a duração das fixações está relacionado à semelhança observada nas amplitudes sacádicas, pois uma fixação de longa duração é precedida por um movimento sacádico de grande amplitude (Cabestrero et al., 2005).

Ademais, a duração da primeira fixação também não foi modulada diferentemente pelos fundos de tela. Este é um tema controverso. Por exemplo, no clássico estudo de Loftus e Mackworth (1978) não houve diferença na duração da primeira fixação nos alvos informativos e não-informativos, diferença esta que foi encontrada em outros trabalhos (De Graef, Christiaens & d'Ydewalle, 1990, Henderson et al, 1999). Contudo, não há informações sobre a duração da primeira fixação em uma tarefa de comparação de tamanho de estímulos

inseridos em gradientes de textura. Dessa maneira, nossos resultados fornecem as primeiras evidências de que as informações de profundidade não influenciam os processos relacionados à primeira fixação, como por exemplo, a aquisição da informação visual (Henderson, 1993; Van Diepen & Wampers, 1998) e, conseqüentemente, a programação da sequência de movimentos oculares (Zingale & Kowler, 1987).

Por outro lado, a análise estatística indicou que os fundos de tela não afetaram diferencialmente o número de fixações realizadas pelos participantes para explorar a imagem. Apesar disso, observou-se uma tendência do gradiente fotografado e do gradiente de perspectiva receberem mais fixações do que a condição controle. Acredita-se que em maiores tempos de exposição, esta diferença seria significativa. Este resultado é consonante com o trabalho de Jansen et al. (2009), que encontraram um aumento no número de fixações à medida que se acrescentavam dicas de profundidade.

Os participantes mostraram uma grande concentração das fixações na região central da imagem. Ao investigar este desvio central, Tatler et al. (2005) não encontraram evidências de que o desvio central possa ser explicado pelo ponto de fixação, que na maioria das vezes está localizado na região central da imagem. Em outro estudo, Tatler (2007) não encontrou evidências de que este fenômeno esteja relacionado a um possível desvio central na saliência visual das imagens. Ou seja, mesmo em cenas com maior saliência nas laterais, os observadores fixam mais o centro da imagem. O mesmo autor advoga que a explicação mais provável é que esta é a melhor estratégia para extrair rapidamente uma representação geral da imagem (*scene gist*). Assim, inicialmente a visão seria orientada para o centro e as fixações posteriores variam de acordo com a tarefa a ser realizada.

O desvio central observado revela outro achado interessante: a grande consistência na localização das primeiras fixações entre os participantes. Parkhurst et al. (2002) defendem que não haveria consistência nas primeiras fixações, caso estratégias *top-down* estivessem envolvidas na seleção da posição destas. Dessa maneira, corroborando o estudo destes autores, os resultados indicam que as primeiras fixações realmente são mediadas por mecanismos *bottom-up*. Ademais, ao analisar a sequência das fixações pode-se afirmar que esta consistência diminui com mais tempo de observação, o que sugere o início do envolvimento de mecanismos *top-down*. Os resultados mostram que a segunda e terceira fixações são preferencialmente dirigidas à região central, embora na terceira já se observe mais fixações em outras regiões. Resultado semelhante foi descrito por Tatler (2007).

De modo bastante interessante, observou-se que a partir da quarta fixação as informações de profundidade modificaram a estratégia visual dos participantes. A quarta e quinta fixações no gradiente de perspectiva e a quinta e sexta fixações no gradiente fotografado estavam localizadas preferencialmente no círculo acima. Por outro lado, na condição controle, a sexta fixação estava localizada com maior frequência no círculo abaixo. Estes resultados revelam a participação da atenção à medida que há mais tempo para explorar a imagem. Além disso, sugerem que a atenção é alocada para o objeto percebido como mais distante (círculo acima) nas condições com gradiente de textura. Parks e Corballis (2006) também encontraram que a atenção é distribuída diferencialmente em profundidade e é direcionada para os estímulos percebidos como mais distantes.

Caso se exclua as fixações dirigidas à área PF, os dados mostram um maior número de fixações no círculo localizado no campo visual superior, independente da ordem das fixações. Cabe ressaltar que esta diferença não se refletiu na duração total média das fixações, ou seja, o tempo de observação do campo visual superior e inferior foi semelhante. Também se observou que o primeiro movimento ocular é dirigido à porção superior da imagem em todos os fundos de tela. Estes resultados são contrários aos estudos de Castelhamo et al. (2009) e Foulsham e Kingstone (2010), os quais apresentam evidências de que os objetos percebidos como próximos recebem mais fixações e que os olhos movem-se para baixo na primeira sacada, respectivamente. Cabe ressaltar que estes trabalhos são condizentes com a hipótese de seleção para ação.

No entanto, os resultados aqui apresentados se coadunam com vários estudos anteriores. Por exemplo, Jansen et al. (2009) encontraram uma preferência em olhar acima nos primeiros 10 s durante a observação de uma imagem bidimensional. Wexler e Ouarti (2008) mostraram que em 82% das vezes, a primeira sacada dirigia-se no sentido do aumento de profundidade. Os autores argumentam que dirigir o olhar inicialmente para uma posição mais distante maximiza a aquisição das informações. Resultados consonantes foram encontrados em um estudo com macaco-rhesus, o qual revelou assimetria nas sacadas verticais: as que são realizadas para cima têm latência e duração menores, maior velocidade e menos erros que as efetuadas para baixo (Zhou & King, 2002).

Os resultados acima descritos são contraditórios às vastas evidências de um melhor desempenho no campo visual inferior? A resposta é não e o trabalho supracitado de Zhou e King (2002) nos fornece dicas valiosas para compreender isto. Estes autores mostraram que

há uma maior sensibilidade atensiva e temporal para estímulos no campo visual superior, sendo importante para a representação do espaço a grandes distâncias, o que é condizente com os nossos achados. Assim, a vantagem do meridiano inferior refere-se à resolução espacial, servindo à representação do espaço próximo ao observador. Inclusive, sabe-se que há uma maior densidade de fotorreceptores na retina superior, a qual representa a parte inferior do espaço visual.

Pode-se questionar ainda que os resultados encontrados são um artefato da localização do estímulo teste, cuja posição sempre foi o campo visual superior. Assim, os participantes notariam que o círculo apresentado no campo visual inferior não mudava de tamanho, direcionando o olhar apenas para a parte superior da imagem, onde o tamanho do estímulo era variável. No entanto, esta crítica não se justifica, pois o mesmo padrão de resultados foi observado em um experimento-piloto realizado anteriormente, no qual o estímulo teste era apresentado ora acima ora abaixo no campo visual.

Além da influência da localização espacial, o tamanho dos estímulos também se mostrou relevante e afetou os parâmetros oculares investigados. Observou-se que o círculo maior recebeu mais fixações e foi inspecionado por mais tempo do que o círculo menor nos três fundos de tela. Por outro lado, o efeito do tamanho dos estímulos sobre a direção do primeiro movimento sacádico foi diferente em função do fundo de tela. Na condição controle, a primeira sacada foi distribuída igualmente entre os círculos maior e menor. Nas condições com informação de profundidade, o círculo maior foi o alvo preferido para a primeira sacada.

Tomados em conjunto, estes resultados sugerem que a atenção tenha sido alocada diferencialmente ao estímulo de tamanho maior e, provavelmente, tanto a atenção sustentada quanto a atenção transitória estão implicadas neste processo. Sabe-se que há um tempo de no mínimo 300 ms entre a apresentação de um estímulo e o recrutamento da atenção sustentada (Nakayama & Mackeben, 1989). Assim, pode-se inferir que ao longo da observação da imagem, os participantes dirigiram seletivamente sua atenção para os estímulos maiores, o que se refletiu na maior quantidade de fixações e na maior duração destas no círculo maior. Em contraste, a evidência de que o primeiro movimento sacádico foi dirigido ao círculo maior sugere o envolvimento da atenção transitória, que é dirigida rápida e automaticamente a objetos no campo visual. Evidências recentes mostram que esta atenção é recrutada por estímulos maiores, tanto em tamanho físico quanto em tamanho percebido (Proulx, 2010; Proulx & Green, 2011). Nossos resultados, porém, indicam a participação da atenção apenas

nas condições com gradiente de textura, mostrando que as informações de profundidade modulam a atenção mesmo em seus estágios iniciais.

5. Discussão geral

O objetivo principal deste estudo foi investigar a dinâmica temporal da representação do tamanho de objetos inseridos em gradientes de textura com informações de profundidade. Estas informações eram provenientes de desenho gerado computacionalmente ou de fotografia com elementos naturais. Segundo Watanabe (2006), estas duas maneiras de representar o espaço são equivalentes e igualmente eficientes para informar sobre a profundidade, porém não há resultados que indiquem se há diferença no processamento destes gradientes em tempos breves de exposição. Estudou-se a percepção de tamanho, pois esta guarda uma estreita relação com a percepção de profundidade e é por ela afetada. Desta maneira, diferenças no tamanho percebido dos estímulos em função das informações de profundidade foram interpretadas como evidência do processamento destas.

Os resultados do primeiro experimento indicaram que as informações de tamanho e profundidade são processadas de maneira muito rápida em nosso sistema visual, ocorrendo em tempos inferiores a 150 ms. Contudo, o tempo mínimo necessário para o processamento inicial e integração destas informações foi diferente entre os gradientes de textura investigados. As maiores distorções de tamanho nas comparações realizadas sob gradiente de perspectiva desenhado foram observadas aos 100 ms, resultado semelhantes aos descritos por Bernardino (2008), Champion e Warren (2008) e Torro-Alves et al. (2012). Em contraste, nas comparações realizadas no gradiente fotografado composto de elementos naturais, a rapidez para extrair as informações de profundidade foi menor, resultando em maiores distorções de tamanho aos 150 ms.

Desta maneira, estudos posteriores que investiguem o processamento inicial de informações de profundidade provenientes de gradientes fotografados devem levar em consideração que, neste caso, se requer mais tempo para apreensão destas informações. Cabe ressaltar que a escassez de estudos que empregam fotografias de cenas ou mesmo de elementos naturais para o estudo da percepção de tamanho e de profundidade contrasta com a importância destes estudos, dada a maior validade ecológica de suas conclusões. Além disso, evidencia-se que é necessária precaução ao extrapolar conclusões de estudos que utilizam gradientes de textura gerados computacionalmente para contextos mais naturais, principalmente no que se refere aos estágios iniciais de processamento.

Ademais, os resultados deste primeiro experimento mostraram uma diminuição na magnitude das distorções de tamanho à medida que o tempo de exposição aumentava,

principalmente após os 100 ms para o gradiente de perspectiva e após os 150 ms para o gradiente fotografado. Provavelmente a atenção está implicada na redução das distorções e, para verificar esta hipótese, empreendeu-se o segundo experimento. Foi realizada uma tarefa semelhante e registraram-se os movimentos oculares, os quais foram utilizados como indicadores da alocação da atenção, dada a intrínseca relação existente entre eles. Nota-se que o segundo experimento não é uma réplica exata do Experimento 1 somado à utilização do *eye tracker*. Optou-se por mudar o delineamento experimental, pois o primeiro e o segundo experimentos não tinham objetivos idênticos, mas sim complementares: investigar mecanismos perceptivos e cognitivos do processamento visual na realização da tarefa, respectivamente.

No primeiro experimento, a instrução indicava ao participante que não movimentasse os olhos, mantendo-os na posição onde é apresentado o ponto de fixação mesmo após o desaparecimento deste. Um problema óbvio desta abordagem é que se deve confiar que os participantes seguiram estritamente esta instrução. Com o uso do *eye tracker* no Experimento 2, pode-se efetuar um maior controle e identificar a ocorrência de movimentos oculares na realização de uma tarefa de comparar tamanhos. Para melhor responder a esta questão, a instrução foi modificada: após o desaparecimento do ponto de fixação, os observadores estavam livres para olhar para onde lhes conviesse. Mesmo com a “permissão” para mover os olhos, isto foi observado com uma frequência muito baixa, sugerindo que mecanismos atentos não foram recrutados ou pelo menos não explicam completamente os achados do Experimento 1.

Os resultados de ambos os experimentos indicam que os participantes foram capazes de processar toda a informação contida nas imagens (círculos e gradiente de textura) de maneira muito rápida, em aproximadamente 100-150 ms, e com apenas uma fixação. Em consonância com estes dados, Palmer, Verghese e Pavel (2000) encontraram que é possível processar múltiplas localizações em paralelo com eficiência semelhante em apresentações breves que não permitem movimentos oculares. Nesta direção, também há indícios de que o processamento de profundidade é organizado em mapas retinotópicos, sendo realizado de maneira automática e em paralelo (Chelazzi, 1999), o que repercute na dificuldade em ignorar estas informações numa tarefa de comparar tamanhos (Farran et al., 2009).

Nosso resultados também são coerentes com os achados de Kirchner e Thorpe (2006), que mostraram que a análise grosseira inicial de uma imagem complexa é realizada em um

tempo não superior a 100 ms. Além disso, há evidências de que uma faixa de tempo semelhante é suficiente tanto para a ativação de neurônios na área V1 que respondem às superfícies, como é o caso dos gradientes de textura (Di Lollo et al., 2000) quanto para atuação de mecanismos de retroalimentação entre as áreas extra-estriadas e a área V1 (Tong, 2003). Ademais, Morvan e Wexler (2005) indicaram que a representação retinal, incluso a de tamanho, é acessada em torno de 100 ms e Cauquil et al. (2006) mostraram que não somos sensíveis às dicas de perspectiva antes de 90 ms. Tomados em conjunto, estes achados revelam que ambas as informações são processadas em tempos semelhantes, próximo aos 100 ms, sem haver prevalência de uma sobre a outra.

Frente a estas evidências, parece não haver dúvidas quanto ao envolvimento de mecanismos *bottom-up* no processamento das informações de tamanho e profundidade. Entretanto, há uma questão ainda controversa no que se refere à representação destas informações: ela ocorre de maneira conjunta e integrada, como predito pela abordagem ecológica proposta por Gibson (1950, 1979) ou a distância (profundidade) é registrada primeiro e infere-se o tamanho a partir desta, como postulado pela Hipótese da Invariância Tamanho-Distância (HITD)? Nossos resultados não permitem chegarmos a uma resposta categórica como Tozawa (2010), cujos resultados dão suporte ao modelo gibsoniano, isto é, as informações de tamanho e de profundidade são processadas rápida e conjuntamente, não havendo a necessidade de mediações cognitivas ou inferenciais para serem representadas no sistema visual.

No segundo experimento também se empreendeu uma análise acerca do padrão dos movimentos oculares no processamento das informações de tamanho e profundidade em uma condição com menor restrição temporal. Os resultados forneceram valiosas evidências sobre o envolvimento de processos *bottom-up* e *top-down* na representação do tamanho e da profundidade em nosso sistema visual.

Primeiramente, cabe salientar que a dimensão espacial das fixações e dos movimentos sacádicos foi modificada pela presença dos gradientes de textura. Observou-se uma preferência em olhar para o círculo apresentado no campo visual superior nas condições com informações de profundidade, o que pode ser interpretado como uma alocação diferencial da atenção para objetos percebidos como mais distantes. Apesar dos resultados não corroborarem os achados de Castelhamo et al. (2009) e Jansen et al. (2009), provavelmente esta estratégia

visual foi empregada pois mudar a atenção de um objeto distante para um objeto próximo é mais rápido do que o contrário, como foi demonstrado por Parks e Corballis (2006).

Os resultados ainda mostraram que a alocação da atenção foi afetada pelo tamanho dos estímulos e pelos gradientes de textura. Observou-se que os estímulos maiores receberam mais fixações tanto nos momentos iniciais de apresentação quanto ao longo da tentativa, apenas nas condições com informação de profundidade. Isto demonstra a participação da atenção transitória e sustentada na aquisição das informações de tamanho. Além disso, os resultados confirmam a pertinência de estudar as mudanças na atenção por meio das fixações, como proposto por Hayhoe e Ballard (2005).

Em contraste, os gradientes de textura não afetaram significativamente a dimensão temporal das fixações e das sacadas. Nos últimos anos, crescentes evidências contrapõem-se à clássica ideia de que a duração de uma fixação é um parâmetro fundamental para compreender o processamento subjacente das informações contidas em uma imagem (Buswell, 1935; Yarbus, 1967).

De acordo com Phillips e Edelman (2008), fixações com duração superior a 200 ms não melhoram o desempenho dos participantes em uma tarefa de discriminação. Assim, os autores defendem que as investigações não devem se ocupar unicamente sobre a duração da fixação, pois esta não reflete o tempo necessário para processar o que é visto durante a fixação; mais importante é compreender o que é possível perceber durante uma fixação. Inclusive, processar de maneira eficiente as informações obtidas em uma fixação é mais relevante do que integrar perfeitamente as informações obtidas em todas as fixações (Najemnik & Geisler, 2005). Tais argumentos estão em acordo com o posicionamento de Vaughan (1982), para quem a duração das fixações é pré-programada, não sendo modificada nem pela saliência dos estímulos (*bottom-up*) e tampouco pelos objetivos do participante na tarefa (*top-down*).

Desta maneira, pesquisas futuras acerca do espaço visualmente percebido em geral e do processamento das informações de tamanho e profundidade em específico devem investigar principalmente a dimensão espacial dos parâmetros oculares em detrimento dos aspectos temporais, porém não se limitar a eles. Hayhoe e Ballard (2005) propõem que mais relevante do que saber onde o participante olhou, é perguntar por que e quando esta região foi selecionada. Nossa análise sobre a sequência das fixações aproximou-se desta abordagem, revelando que a seleção inicial da localização das fixações pode ser atribuída a processos

bottom-up e o momento em que os mecanismos *top-down* começaram a afetar a estratégia visual dos participantes em função da presença ou não de informações de profundidade. Ademais, estudos ulteriores devem considerar que os parâmetros oculares modificam-se não apenas pelo tipo de tarefa realizada, mas também pela instrução dada aos participantes (Watson et al., 2010).

A investigação empreendida no segundo experimento com o tempo de 1500 ms foi uma abordagem inicial e exploratória sobre a participação dos mecanismos *bottom-up* e *top-down* no processamento das informações de tamanho e profundidade. Novas investigações são necessárias para se alcançar uma compreensão completa acerca destes processos, o que requiere conhecer as repercussões de um sobre o outro e como se influenciam mutuamente. Também é imprescindível investigar o envolvimento *top-down* na representação do espaço, bem como o efeito de diferentes processos atentos, tais como a atenção transitória e sustentada, na percepção visual de tamanho.

Não se pode negligenciar que o uso de técnicas de imageamento cerebral é uma poderosa ferramenta, que também deve ser utilizada para o estudo da percepção de tamanho. Em um elegante estudo, Schwarzkopf, Song e Rees (2010) mostraram haver uma correlação negativa entre o tamanho da área visual primária (V1) e a magnitude das ilusões de Ebbinghaus e Ponzo. Isto é, menor a área de V1, mais forte é a ilusão percebida pelo participante. Com este resultado, pode-se concluir que a experiência consciente está fortemente relacionada à estrutura cortical, o que consequentemente denota o seu caráter idiossincrático. O trabalho de Schwarzkopf et al. (2010) ainda chama a atenção para uma relação amiúde negligenciada entre consciência e percepção visual.

Pode-se concluir, portanto, que o estudo da percepção de tamanho não é trivial, pois ajuda-nos a compreender o funcionamento da consciência. Esta é tarefa árdua e uma das fronteiras finais no estudo do cérebro humano. A dificuldade reside no fato de que se utiliza a própria consciência para investigá-la. É como estar em um barco em alto mar e querer desmontá-lo para conhecer seu funcionamento. Contudo, da dificuldade e complexidade não se segue impossibilidade. Esforços têm sido empreendidos e resultados fascinantes aparecem, inclusive com repercussões na saúde pública, como é o caso do interessante trabalho de Simas et al. (2011). Os resultados mostraram que pacientes com esquizofrenia, cujos processos conscientes estão severamente prejudicados, apresentam alterações na percepção de tamanho

e de forma. Sugere-se, inclusive, que tais alterações podem ser utilizadas no diagnóstico precoce dos sintomas positivos da doença.

Enfim, uma interessante e ainda pouco explorada abordagem para o estudo da dinâmica temporal dos processos perceptivos é o modelo Bayesiano. Nesta perspectiva, a visão é um processo inferencial e sequencial de tomada de decisão, baseada em probabilidades condicionais, que são constante e continuamente atualizadas. Segundo Hegdé (2008), este modelo é completo e robusto o suficiente para abarcar e estender a abordagem hierárquica clássica (*coarse-to-fine*), que divide os processos em *bottom-up* e *top-down*. Contudo, o modelo Bayesiano não foi adequada e extensivamente testado, tarefa a ser executada nos próximos anos.

6. Conclusão

O estudo apresentado permitiu um maior conhecimento sobre os mecanismos perceptuais e atentos envolvidos no processamento das informações de tamanho e de profundidade. Os resultados mostraram que o processamento de informações de profundidade geradas no computador é temporalmente diferente das provenientes de fotografias. Apesar desta diferença, é surpreendente a velocidade com que apreendemos informações complexas com apenas uma fixação e construímos uma percepção do espaço, se não completa, suficientemente acurada, que nos permite tomar decisões e executar ações em resposta às demandas ambientais. Por fim, revelou-se que a presença das informações de profundidade afetou a alocação inicial e tardia da atenção em função do tamanho dos estímulos e, além, disso, modificou a dimensão espacial dos parâmetros oculares, mas não a sua dimensão temporal.

Referências¹

- Aks, D. J., & Enns, J. T. (1996). Visual search for size is influenced by a background texture gradient. *Journal of Experimental Psychology: Human Perception and Performance*, 22(6), 1467-1481.
- Allison, R. S., & Howard, I. P. (2000). Temporal dependencies in resolving monocular and binocular cue conflict in slant perception. *Vision Research*, 40 (4), 1869-1885.
- Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, 103, 62-70.
- Anton-Erxleben, K., Henrich, C., & Treue, S. (2007). Attention changes perceived size of moving visual pattern. *Journal of Vision*, 7, (11), 1-9.
- Awh, E., Armstrong, K. M., & Moore, T. (2006). Visual and oculomotor selection: links, causes and implications for spatial attention. *Trends in Cognitive Science*, 10, 124-130.
- Bengston, J. K., Stergios, J. C., Ward, J. L., & Jester, R. E. (1980). Optic array determinants of apparent distance and size in pictures. *Journal of Experimental Psychology: Human, perception and performance*, 6(4), 751-759.
- Bernardino, L. G. (2008). *Comparações de tamanhos sob informações pictóricas de profundidade provenientes de gradientes de textura e horizonte em exposições breves*. Dissertação de Mestrado, Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo, Ribeirão Preto.
- Bernardino, L. G., Sousa, B. M., Fukusima, S. S. (2010). Psychophysical Method Comparison To Estimate Psychometric Functions For Visually Perceived Sizes. In *26th Annual Meeting of the International Society for Psychophysics - Fechner Day*, (pp. 311-314 2010), Padova, Italy.
- Bian Z., Andersen G. J. (2010). The advantage of a ground surface in the representation of visual scenes. *Journal of Vision*, 10(8), 1-19.

¹ De acordo com a *American Psychological Association (APA)*

- Biederman, I., Rabinowitz, J. C., Glass, A. L. & Stacey, E. W. J. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, 103, 597-600.
- Bourne, V. J. (2006). The divided visual field paradigm: methodological considerations. *Laterality*, 11, 373-393.
- Bruggemann, H., Yonas, A., & Konczak, J. (2007). Processing linear perspective information for perception and action: The effects of binocular and monocular presentation. *Neuropsychologia*, 45(7), 1420-1426.
- Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.
- Cabestrero, R., Conde-Guzón, P.A., Crespo, A., Grzib, G., & Quirós, P. (2005). *Fundamentos psicológicos de la actividad cardiovascular y oculomotora*. Madrid: UNED ediciones.
- Callaway, E. M. (2005). Structure and function of parallel pathways in the primate early visual system. *Journal of Physiology*, 566, 13-19.
- Cameron, E. L., Tai, J. C., & Carrasco, M. (2002). Covert attention affects the psychometric function of contrast sensitivity. *Vision Research*, 42, 949-967.
- Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, 7, 308-313.
- Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the CSF: support for signal enhancement. *Vision Research*, 40, 1203-1215.
- Castelhano, M.S., Mack, M., & Henderson, J.M. (2009). Viewing task influences eye movements during active scene perception. *Journal of Vision*, 9(3)6, 1-15.
- Cauquil A. S., Trotter, Y., & Taylor M. J. (2006). At what stage of neural processing do perspective depth cues make a difference? *Experimental Brain Research*, 170(4) 457-463.

- Champion, R. A., & Warren, P. A. (2008). Rapid size scaling in visual search. *Visual Research*, 48, 1820-1830.
- Chelazzi, L. (1999). Serial attention mechanisms in visual search: A critical look at the evidence. *Psychological Research*, 62(2-3), 195-219.
- Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A. et al. (1998). A common network of functional areas for attention and eye movements. *Neuron*, 21, 761-773.
- Cornsweet, T. (1962). The staircase method in psychophysics. *American Journal of Psychology*, 75, 485-491.
- Cristino, F., & Baddeley, R. (2009). The nature of visual representations involved in eye movements when walking down the street. *Visual Cognition*, 17, 880-903.
- Cutting, J. E., & Millard, R. T. (1984). Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology: General*, 113(2), 198-216.
- Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.), *Handbook of perception and cognition: Vol. 5. Perception of space and motion* (pp. 69-117). San Diego: Academic Press.
- Day, R. H., & Parks, T. E. (1989). To exorcise a ghost from the perceptual machine. In M. Hershenov (Ed.), *The moon illusion* (pp. 343-350). Hillsdale, NJ: Erlbaum.
- De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, 52, 317-329.
- Di Lollo, V., Enns, J. T., Resink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, 129, 481-507.
- Dosher, B. A., & Lu, Z. -L. (2000). Mechanisms of perceptual attention in precuing of location. *Vision Research*, 40(10-12), 1269-1292.

- Epstein, W. (1966). Perceived depth as a function of relative height under three background conditions. *Journal of Experimental Psychology*, 72, 335-338.
- Epstein, W., Park, J., & Casey, A. (1961). The current status of the size-distance hypotheses. *Psychological Bulletin*, 58 (6), 491-514.
- Farran, E. K., Whitaker, A., & Patel, N. (2009). The effect of pictorial depth information on projected size judgments. *Attention, Perception & Psychophysics*, 71(1), 207-214.
- Foley, J. M., Ribeiro-Filho, N. P., & Da Silva, J. A. (2004). Visual perception of extent and the geometry of visual space. *Vision Research*, 44 (2), 147-156.
- Foulsham, T., & Kingstone, A. (2010). Asymmetries in the direction of saccades during perception of scenes and fractals: Effects of image type and image features. *Vision Research*, 50 (8), 779-795.
- Found, A., & Müller, H. J. (2001). Efficient search for size targets on a background texture gradient: Is detection guided by discontinuities in the retinal-size gradient of items? *Perception*, 30(1), 21-48.
- Frichtel, M., & Lécuyer, R. (2007). The use of perspective as a depth cue with a 2D display in 4 and 5-month-old infants. *Infant Behavior and Development*, 30, 409-421.
- Gareze, L., & Findlay, J. M. (2007). Absence of scene context effects in object detection and eye gaze capture. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds), *Eye movements: A window on Mind and Brain* (pp 617-637). Oxford: Elsevier.
- Gazzaniga, M.S., & Ladavas, E. (1987). Disturbances in spatial attention following lesion of the right parietal lobe. In: M. Jeannerod (Ed.), *Neurophysiological and Neuropsychological Aspects of Spatial Neglect* (pp. 203-213). Amsterdam: Elsevier Science Publishers.
- Gibson, J. J. (1950). *The perception of the visual world*. New York: Houghton Mifflin.
- Gibson, J. J. (1979) *The ecological approach to visual perception*. Boston: Houghton Mifflin.

- Gibson, J. J., & Bridgeman, B. (1987). The visual perception of surface texture in photographs. *Psychological Research*, 49(1), 1-5.
- Gilbert, C. D., Sigman, M., & Crist, R.E. (2001) The neural basis of perceptual learning. *Neuron*, 31, 681-697.
- Gogel, W. C., & Da Silva, J. A. (1987). A two-process theory of the response to size and distance. *Perception & Psychophysics*, 41(3), 220-238.
- Goldstein, E. B. (2009). *Sensation and perception* (8th ed.). Belmont, CA: Cengage Learning Int.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20-25.
- Goodenough, B., & Gilliam, B. (1997). Gradients as visual primitives. *Journal of Experimental Psychology: Human Perception and Performance*, 23(2), 370-387.
- Green, D. M., & Swets, J. A. (1966/1988). *Signal detection theory and psychophysics*, reprint edition. Los Altos, CA: Peninsula Publishing.
- Gruber, H. E. (1954). The relation of perceived size to perceived distance. *American Journal of Psychology*, 67, 411-426.
- Gulick, W. L., & Stake, R. E. (1957). The effect of time on size-constancy. *American Journal of Psychology*, 70(2), 276-279.
- Haber R. N., & Levin C. A. (2001). The independence of size perception and distance perception. *Perception & Psychophysics*, 63(7), 1140-1152.
- Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9(4), 188-193.
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of awareness. *Nature*, 383, 334-338.

- He, Z. J., & Nakayama, K. (1995). Visual attention to surfaces in 3-D space. *Proceedings of the National Academy of Sciences of the USA*, 92, 11155 – 11159.
- Hecht, H., van Doorn, A., & Koenderink, J. J. (1999). Compression of visual space in natural scenes and in their photographic counterparts. *Perception & Psychophysics*, 61(7), 1269-1286.
- Hegd, J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology*, 84, 405-439.
- Hemker, L., & Kavek, M. (2010). The relative contribution of relative height, linear perspective, and texture gradients to pictorial depth perception in 7-month-old infants. *Perception*, 39, 1476-1490.
- Henderson, J. M. (1992). Visual attention and eye movement control during reading and picture viewing. In K. Rayner. *Eye Movements and Visual Cognition* (pp. 269-283). New York: Springer-Verlag.
- Henderson, J. M. (1993). Visual attention and saccadic eye movements. In G. d'Ydewalle & J. Van Rensbergen (Eds.), *Perception and Cognition* (pp. 37-50). Amsterdam: Elsevier.
- Henderson, J. M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 269-283). Oxford: Elsevier
- Henderson, J. M., Weeks, P. A., Jr., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during scene viewing. *Journal of Experimental Psychology: Human Perception & Performance*, 25, 210-228.
- Higashiyama, A. (1979). The perception of size and distance under monocular observation. *Perception & Psychophysics*, 26(3), 230-234.
- Higashiyama, A., & Shimono, K. (2004). Mirror vision: Perceived size and perceived distance of virtual images. *Perception & Psychophysics*, 66, 679-691.
- Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, 4(12), 967-992.

- Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, 57, 787-795.
- Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology*, 54, 21-37.
- Hooze I. Th. C., & Erkelens, C. J. (1996). Control of fixation duration in a simple visual search task. *Perception and Psychophysics*, 58(7), 969-976.
- Howell, D. C. (2002). *Statistical Methods for Psychology*. Boston: PWS-Kent.
- Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human, Learning and Memory*, 6(1), 1-12.
- Jansen, L., Onat, S., & König, P. (2009). Influence of disparity on fixation and saccades in free viewing of natural scenes. *Journal of Vision*, 9(1)29, 1-19.
- Johnson, J. S., & Olshausen, B. A., (2005). The earliest EEG signatures of object recognition in a cued-target task are postsensory. *Journal of Vision*, 5, 299-312.
- Kandel E. R., Schwartz J. H., & Jessell, T. M. (2000). *Principles of Neural Science* (4th ed.). New York: McGraw-Hill.
- Kessels, R. P. C., Postma, A. & De Haan, E. H. F. (1999). P and M channel-specific interference in the What and Where pathway. *Neuroreport*, 10, 1-3.
- Kilpatrick, F. P., & Ittelson, W. H. (1953). The size-distance invariance hypothesis. *Psychology Review*, 60, 223-231.
- Kirchner H., & Thorpe S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, 46, 1762-1776.
- Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research*, 43, 2539–2558.

- Konen, C., & Kastner, S. (2008) Two hierarchically organized neural systems for object information in human visual cortex. *Nature Neuroscience*, 11(2), 224-231.
- Kraft, R. N., Patterson, J. F., & Mitchell, N. B. (1986). Distance perception in photographic displays of natural settings. *Perceptual and Motor Skills*, 62, 179-186
- Leek, M. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, 63(8), 1279-1292.
- Leigh, R. J., & Zee, D. S. (Eds.). (1991). *The neurology of eyes movement*. Philadelphia: Davis.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2), 467-477.
- Ling, S., & Carrasco, M. (2006). Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision Research*, 46, 1210-1220.
- Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 565-572.
- Loomis J. M., Philbeck J. W., & Zahorik, P. (2002). Dissociation between location and shape in visual space. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1202-1212.
- Lu, Z., & Doshier, B. A. (1998). External noise distinguishes mechanisms of attention. *Vision Research*, 38, 1183-1198.
- Macevoy, S. P., & Fitzpatrick, D. (2006). Visual physiology: perceived size looms large. *Current Biology*, 16(9), 330-332.
- Masin, S. C. (2003). Test of the effect of attention on judged length of a line. *Perceptual and Motor Skills*, 96, 599-606.

- Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, 6, 414-417.
- Miura, T. (1990). Active function of eye movement and useful field of view in a realistic setting. In R. Groner, G. d'Ydewalle and R. Parnham (eds) *From Eye to Mind: Information Acquisition in Perception, Search and Reading* (pp. 119-127). Amsterdam: Elsevier.
- Mon-Williams, M., & Tresilian, J. R. (1999). The size-distance paradox is a cognitive phenomenon. *Experimental Brain Research*, 126(4), 578-582.
- Morita, H., & Kumada, T. (2003). Effects of pictorially-defined surfaces on visual search. *Vision Research*, 43, 1869-1877.
- Morvan, C., & Wexler, M. (2005). Reference frames in early motion detection. *Journal of Vision*, 5(2), 131-138.
- Murray, S.O., Boyaci, H. & Kersten, D. (2006) The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience*, 9, 429-434.
- Najemnik, J., & Geisler, W. S. (2005) Optimal eye movement strategies in visual search. *Nature*, 434, 387-391.
- Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, 29, 1631-1646.
- Navon, D. (1977). Forest before trees: the precedence of global features in visual perception. *Cognitive Psychology*, 9, 353-383.
- Norman, J. (1980). Direct and indirect perception of size. *Perception and Psychophysics*, 28, 306-314.
- Norman, J. (2002). Two visual systems and two theories of perception: an attempt to reconcile the constructivist and ecological approaches. *Behavioral and Brain Science*, 25, 73-144.

- O'Brien, J. M. D. (1997). *Texture and motion cues in slant perception*. PhD. Thesis, London University, London, UK.
- Oruç, I., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision Research*, 43, 2451-2468.
- Ozkan, K., & Braunstein, M. L. (2009) Predominance of ground over ceiling surfaces in binocular rivalry. *Attention, Perception & Psychophysics*, 71(6), 1305-1312.
- Ozkan, K., & Braunstein, M. L. (2010). Background surface and horizon effects in the perception of relative size and distance. *Visual Cognition*, 18(2), 229-254.
- Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge: MIT Press.
- Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, 40, 1227-1268.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42, 107–123.
- Parkhurst, D. J., & Niebur, E. (2003). Scene content selected by active vision. *Spatial Vision*, 16, 125–154.
- Parks, N. A., & Corballis, P. M. (2006). Attending to depth: electrophysiological evidence for a viewer-centered asymmetry. *Neuroreport*, 17(6), 643-647.
- Phillips, M. H., & Edelman, J. A. (2008). The dependence of visual scanning performance on saccade, fixation, and perceptual metrics. *Vision Research*, 48, 926-936.
- Poom, L., & Börjesson, E. (1999). Perceptual depth synthesis in the visual system as revealed by selective adaptation. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1–14.
- Portin, K., Vanni, S., Virsu, V., & Hari, R. (1999). Stronger occipital cortical activation to lower than upper visual field – Neuromagnetic recordings. *Experimental brain research*, 124(3), 287-294.

- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.
- Previc, F. H. (1990). Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications. *Behavioral and Brain Sciences*, 13, 519-565.
- Prinzmetal W., Leonhardt J., & Garrett R. (2008). Does gaze direction affect accuracy? *Visual Cognition*, 16, 567-584.
- Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: Reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General*, 134, 73-92.
- Prinzmetal, W., & Wilson, A. (1997). The effect of attention on phenomenal length. *Perception*, 26, 193-205.
- Proulx, M. J. (2010). Size Matters: Large Objects Capture Attention in Visual Search. *PLoS ONE*, 5(12), e15293.
- Proulx, M. J. & Green, M. (2011). Does apparent size capture attention in visual search? Evidence from the Müller-Lyer illusion. *Journal of Vision*, 11(13)21, 1-6.
- Qu, Z., Song, Y., & Ding, Y. L. (2006). Asymmetry between the upper and lower visual fields: An event-related potential study. *Chinese Science Bulletin*, 51(5), 536-541.
- Rao, S. C., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, 276, 821-824.
- Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2004a). Processing of one, two or four natural scenes in humans: the limits of parallelism. *Vision Research*, 44, 877-894.
- Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2004b). How parallel is visual processing in the ventral pathway? *Trends in cognitive sciences*, 8(8), 363-370.
- Reynolds, R.I. (1981). Perception of an illusory contour as a function of processing time. *Perception*, 10, 107-115.

- Richards, W., & Miller, J. F. Jr. (1971). The corridor illusion. *Perception & Psychophysics*, 9, 421-423.
- Riggs, L. A., Ratliff, F., Cornsweet, J. C., & Cornsweet, T. N. (1953). The disappearance of steadily fixated visual test objects. *Journal of the Optical Society of America*, 43, 495-501.
- Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, 25, 31-40.
- Ruch, T. C., & Fulton, J. F. (1960) *Medical physiology and biophysics*. Philadelphia, PA: W. B. Saunders Company.
- Salthouse, T. A., & Ellis, C. L. (1980). Determinants of eye-fixation duration. *American Journal of Psychology*, 93, 207-234.
- Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye tracking protocols. *Proceedings of the 2000 symposium on Eye tracking research & application* (pp. 71-78). Palm Beach Gardens, FL, USA.
- Santos, N. A., Nogueira, R. M. T. B. L., & Simas, M. L. B. (2005). Processamento visual da forma: evidências para canais múltiplos de frequências angulares em humanos. *Psicologia: Reflexão e Crítica*, 18(1), 98-103.
- Saunders, J. A., & Backus, B. T. (2006). Perception of surface slant from oriented textures. *Journal of Vision*, 6(9), 882-897.
- Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature Neuroscience*, 14, 28-30.
- Schiffman, H. R. (2005). *Sensação e Percepção*. Rio de Janeiro: LTC.
- Segev, R., Schneidman, E., Goodhouse, J., & Berry, M. J. 2nd. (2007) Role of eye movements in the retinal code for a size discrimination task. *Journal of Neurophysiology*, 98, 1380-1391.

- Simas, M. L. B., Nogueira, R. M. T. B. L., Menezes, G. M. M., Amaral, V. F., Lacerda, A. M., & Santos, N. A. (2011) O uso de pinturas de Dalí como ferramenta para avaliação das alterações na percepção de forma e tamanho em pacientes esquizofrênicos. *Psicologia USP*, 22(1), 67-80.
- Sinai, M. J., Ooi T. L., & He, Z. J. (1998). Terrain influences the accurate judgement of distance. *Nature*, 395(6701), 497-500.
- Smith, O. W. (1958). Judgments of size and distance in photographs. *The American Journal of Psychology*, 71(3), 529-538.
- Sperandio, I., Savazzi, S., Gregory, R. L., Marzi, C. A. (2009). Visual reaction time and size constancy. *Perception*, 38, 1601-1609.
- Spivey, M. J., & Spirn, M. J. (2000). Selective visual attention modulates the direct tilt aftereffect. *Perception and Psychophysics*, 62(8), 1525-1533.
- Talgar, C. P., & Carrasco, M. (2002). Vertical meridian asymmetry in spatial resolution: visual and attentional factors. *Psychonomic Bulletin & Review*, 9(4), 714-722.
- Tatler, B. W. (2007). The central fixation bias in scene viewing: selection an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, 7, 1-17.
- Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005). Visual correlates of fixation selection: effects of scale and time. *Vision Research*, 45(5), 643-659.
- Tatler, B., Wade, N., Kwan, H., Findlay, J., & Velichkovsky, B. (2010). Yarbus, eye movements, and vision. *i-Perception*, 1, 7-27.
- Theeuwes, J., Kramer, A. F., & Kingstone, A. (2004). Attentional capture modulates perceptual sensitivity. *Psychonomic Bulletin & Review*, 11(3), 551-554.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520-522.

- Todd, J. T., & Akerstrom, R. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance*, 13(2), 242-55.
- Tong, F. (2003). Primary visual cortex and visual awareness. *Nature Review: Neuroscience*, 4, 219-229.
- Torro-Alves, N., Bacci, A. V. F., Bernardino, L. G., & Fukusima, S. S. (2012). Efeitos da perspectiva na amplificação do tamanho de objetos que interceptam a linha do horizonte. *Psicologia: Reflexão e Crítica*, 25(1), 182-187.
- Torro-Alves, N., & Fukusima, S. S. (2007). Gradientes de textura com linhas de perspectiva e horizonte amplificam a superestimação de tamanho relativo em espaços pictóricos. *Psicologia: Reflexão e Crítica*, 20 (1), 35-42.
- Tozawa, J. (2010). Role of texture gradient in the perception of relative size. *Perception*, 39 (5), 641-660.
- Tozawa, J., & Oyama, T. (2006). Effects of motion parallax and perspective cues on perceived size and distance. *Perception*, 35(8), 1007-1023.
- Tsal, Y., Shalev, L., & Zakay, D. (2005). The lengthening effect revisited: a reply to Prinzmetal and Wilson (1997) and Masin (1999). *Psychonomic Bulletin & Review*, 12 (1), 185-190.
- Tsutsui, K., Jiang, M., Yara, K., Sakata, H., & Taira, M. (2001). Integration of perspective and disparity cues in surface-orientation-selective neurons of área CIP. *Journal of Neurophysiology*, 86, 2856-2867.
- Tsutsui, K.-I., Sakata, H., Naganuma, T. (2002). Neural correlates for perception of 3D surface orientation from texture gradient. *Science*, 298, 409-412.
- van Diepen, P. M. J., De Graef, P., & d'Ydewalle, G. (1995). Chronometry of foveal information extraction during scene perception. *Studies in Visual Information Processing*, 6, 349-362.

- Vaughan, J. (1982). Control of fixation duration in visual search and memory search: Another look. *Journal of Experimental Psychology: Human Perception & Performance*, 8, 709–723.
- Viviani, P. (1990). Eye movements in visual search: Cognitive, perceptual, and motor control aspects. In E. Kowler (Ed.), *Reviews of oculomotor research: Vol. 4. Eye movements and their role in visual and cognitive processes* (pp. 353–393). Amsterdam: Elsevier
- Watanabe, T. (2006). Geometrical structures of photographic and stereoscopic spaces. *Spanish Journal of Psychology*, 9(2), 263-272.
- Watson, M. R., Brennan, A. A., Kingstone, A., & Enns, J. T. (2010). Looking versus seeing: Strategies alter eye movements during visual search. *Psychonomic Bulletin & Review*, 17(4), 543-549.
- Watson, M.R., & Enns, J.T. (2012). Depth perception. In V. S. Ramachandran (Ed.), *Encyclopedia of Human Behavior*, (2nd ed.). (pp. 690-696). San Diego, CA: Elsevier.
- Welchman, A. E., Deubelius, A., Conrad, V., Bühlhoff, H. H., & Kourtzi, Z. (2005). 3D shape perception from combined depth cues in human visual cortex. *Nature Neuroscience*, 8, 820-827.
- Wexler, M., & Ouarti, N. (2008). Depth affects where we look. *Current Biology*, 18, 1868-1872.
- Williams, P. A., & Enns, J. T. (1996). Pictorial depth and framing have independent effects on the horizontal-vertical illusion. *Perception*, 25(8), 921- 926.
- Wu, B., He, Z. J., & Ooi, T. L. (2007). The linear perspective information in ground surface representation and distance judgment. *Perception & Psychophysics*, 69(5), 654-672.
- Yarbus, A. L. (1967). *Eye Movements and Vision*. New York: Plenum Press.
- Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research*, 39, 293-305.

Zingale, C. M., & Kowler, E. (1987). Planning sequences of saccades. *Vision Research*, 27, 1327-1341.

Zhou, W., & King, W. M. (2002). Attentional sensitivity and asymmetries of vertical saccade generation in monkey. *Vision research*, 42(6), 771-779.

Apêndices

APÊNDICE A - Artigo apresentado no exame de qualificação (de acordo com as instruções aos autores da revista "Psicologia: Reflexão e Crítica")

Área na qual se enquadra a submissão: processos básicos.

"A integração tamanho-distância é temporalmente diferente em espaços pictóricos e fotografados"

"Size-distance integration is temporally different in pictorial and photographed space representations"

Bernardino, L. G.¹, Fukusima, S. S.¹

¹ Universidade de São Paulo, Ribeirão Preto, Brasil

Endereço para correspondência:

Leonardo Gomes Bernardino

FFCLRP - Universidade de São Paulo
Programa de Pós-Graduação em Psicobiologia
Av. Bandeirantes, 3900. Monte Alegre.
CEP: 14040-901
Ribeirão Preto - SP - Brasil
FAX: + 55-16-3633-2660
TEL: + 55-16-3602-4448
E-mail: leolokigb@yahoo.com.br

Resumo

O presente estudo teve por objetivo investigar aspectos temporais da integração tamanho-distância em gradientes de textura pictórico e fotografado. Os participantes compararam o tamanho de dois círculos apresentados brevemente no meridiano vertical sobre 3 fundos de tela. Os resultados mostraram maiores distorções de tamanho na condição pictórica (100 ms) em comparação com a condição fotografada (150 ms). Observou-se também uma maior sensibilidade para discriminar tamanhos à medida que o tempo de exposição aumentou. Conclui-se que o tempo necessário para a integração tamanho-distância é diferente em função da representação do espaço (pictórica ou fotografada), embora seja rápida para ambas, dando suporte às hipóteses *bottom-up* de processamento da informação.

Palavras-chave: integração tamanho-distância; gradiente de textura; fotografia; tempo de exposição.

Abstract

The present study aimed to investigate temporal features of size-distance integration on pictorial and photographed gradients. The participants compared two circles presented briefly in the vertical meridian on 3 backgrounds conditions. The results showed greater size distortions in the pictorial condition (100 ms) compared with the photographed condition (150 ms). There was also a greater sensitivity to discriminate sizes as stimuli exposure time increased. We concluded that the time required to size-distance integration is different based on the space representation (pictorial or photographed) and is very quickly in both, supporting *bottom-up* information processing hypothesis.

Keywords: size-distance integration; texture gradient; photograph; exposure time.

Introdução

O tempo todo o sistema visual humano realiza estimativas de tamanho e de distância, o que permite interagir com objetos e pessoas ao redor, tomar decisões, planejar e executar ações. A imagem formada na retina fornece informações sobre tamanho e distância. Quanto mais distante um objeto, menor será o tamanho projetado deste na retina, isto é, tamanho e distância são inversamente proporcionais. Contudo, essa relação não é tão direta e inequívoca como pode parecer. O tamanho percebido altera-se não linearmente (muda menos) em relação ao tamanho da imagem projetado na retina à medida que aumenta a distância do observador (Foley, Ribeiro-Filho & Da Silva, 2004).

Uma teoria amplamente aceita para explicar a relação entre a percepção de tamanho e a percepção de distância é a Hipótese da Invariância Tamanho-Distância, a qual postula que o ângulo visual é transformado em tamanho percebido após a distância ser registrada (Kilpatrick & Ittelson, 1953). Assim, a percepção de tamanho é derivada da percepção de distância. Outra teoria clássica é a proposta por Gibson (1950). Este autor afirma que o gradiente de textura é uma informação suficiente para o cálculo da distância e, através desta, o tamanho é determinado diretamente. Posteriormente, propôs-se que era importante fazer uma distinção entre fatores perceptuais e cognitivos no processamento de tamanho e distância (Gogel & Da Silva, 1987).

Nos últimos anos cresceram as evidências de que o processamento de informações de tamanho requer conhecimento sobre a distância percebida do objeto, porém o contrário não é verdadeiro (Haber & Levin, 2001; Tozawa & Oyama, 2006). Desse modo, a percepção de tamanho e a percepção de distância seriam processos independentes.

Um importante argumento a favor desta dissociação é apresentado por O'Brien (1997). Este autor afirma que se os julgamentos de tamanho ou de distância forem efetuados sem

limitações visuais, os processamentos destas informações não estariam relacionados. Desse modo, se há algum tipo de limitação visual, por exemplo em condição visual monocular, o sistema visual necessita de uma mediação cognitiva, valendo-se da relação entre tamanho e distância para forjar uma percepção mais acurada do espaço.

Na década de 90, dois importantes estudos investigaram a influência dos gradientes de textura sobre a percepção de tamanho. Williams e Enns (1996) mostraram que as distorções de tamanho na ilusão horizontal-vertical são fortemente afetadas pelos gradientes em que estão inseridos os estímulos. Em outro estudo mostrou-se que em uma tarefa de busca visual, o tempo de reação era maior para situações incongruentes entre tamanho e distância, informada por gradiente de textura. Assim, localizar o estímulo-alvo era mais difícil, por exemplo, se este fosse grande e percebido como próximo ao observador (Aks & Enns, 1996).

Uma década depois, Torro-Alves e Fukusima (2007) encontraram uma superestimação de até 10% do tamanho dos estímulos apresentados no campo visual superior, percebidos como mais distantes devido ao gradiente de perspectiva. É importante observar que os estudos supracitados investigaram a integração das informações de tamanho e dos gradientes de textura com estímulos pictóricos, ou seja, desenhos gerados computacionalmente (Aks & Enns, 1996; Williams & Enns, 1996; Torro-Alves & Fukusima, 2007).

Uma interessante perspectiva de investigação deste assunto é o uso de fotografias de cenas reais ou mesmo contendo elementos naturais. Segundo Watanabe (2006), as representações do espaço fotografadas são equivalentes às pictóricas. Para este autor, ambas as representações encerram propriedades não-euclidianas do espaço visualmente percebido.

Além disso, estudos sugerem que a percepção de tamanho e a percepção de distância são afetadas da mesma maneira em fotografias e no ambiente real (Bengston, Stergios, Ward & Jester, 1980) e que as distorções do espaço observadas em ambos são similares (Hecht, van

Doorn & Koenderink, 1999). Assim, o uso de fotografias pode fornecer várias contribuições ao estudo da percepção visual do espaço, dada sua maior validade ecológica.

Também há evidências de que o tamanho é processado de maneira rápida, pois a representação do tamanho é sensível ao tamanho aparente dos estímulos e não apenas ao tamanho retiniano (Found & Muller, 2001). Estes achados psicofísicos são corroborados por estudos neurofisiológicos, que indicam que a atividade de V1 reflete não o tamanho do *input* retiniano, mas sim o tamanho percebido (Murray, Boyaci & Kersten, 2006).

Apesar destas evidências de que a percepção de tamanho é forjada em tempos curtos, há poucos estudos que investigaram diretamente o papel modulador do tempo de exposição no tamanho relativo. Tampouco se sabe se há diferenças temporais no processamento de espaços pictóricos ou fotografados. Desse modo, o presente estudo teve por objetivo investigar o tempo necessário para ocorrer a integração tamanho-distância em gradientes de textura pictórico e fotografado. Com isso, espera-se alcançar um maior conhecimento sobre os mecanismos perceptuais e cognitivos envolvidos no processamento em conjunto de informações de tamanho e de distância.

Método

Participantes

Noventa e seis estudantes (sessenta mulheres) da Universidade de São Paulo (USP), com idade média de 24 anos, participaram do experimento. Todos tinham acuidade visual normal ou corrigida em ambos os olhos. Este estudo foi aprovado pelo Comitê de Ética em pesquisa da Faculdade de Filosofia, Ciências e Letras da USP, conforme as normas vigentes no Brasil sobre experimentos com humanos. Todos os participantes foram voluntários e nenhum recebeu pagamento.

Material e equipamento

Um monitor LG Flatron ez T910B Plus (19”), sob resolução de 1024 x 768 *pixels* e frequência de atualização de 85Hz e um computador com um programa desenvolvido em Lazarus – Free Pascal 0.9.28.2 foram utilizados para apresentar os estímulos e coletar as respostas durante a execução do experimento numa sala escura. O participante, com o olho esquerdo vendado, sentava-se em uma cadeira a uma distância de 57 cm do monitor com sua cabeça posicionada em um apoiador de queixo.

Os estímulos eram dois círculos negros apresentados simultaneamente: um o padrão, apresentado no campo visual inferior com tamanho constante (100 pixels ou 3,56° de ângulo visual), e outro o teste, apresentada no campo visual superior cujo diâmetro variou de 85% a 115% em relação ao estímulo padrão (3,03° e 4,10° de ângulo visual, respectivamente), com diferença de 0,11° entre eles, totalizando 11 estímulos teste.

Estes pares de estímulos foram apresentados no centro do eixo vertical sobre três fundos de tela de cor cinza, como pode ser observado na Figura 1: controle, sem textura (1a); gradiente de perspectiva pictórico (1b); e gradiente de textura fotografado (1c). A fotografia utilizada como gradiente de textura com elementos foi manipulada com o auxílio do programa CorelDRAW X3, para que não houvesse diferença de brilho entre esta e os demais fundos. Em cada condição de fundo de tela, a exposição dos estímulos durava 50 ou 100 ou 150 ou 200 milissegundos (ms).

- Inserir Figura 1 -

Procedimento

Cada tentativa iniciava com a apresentação de um ponto de fixação no centro da tela por 500ms, seguido pelos pares de estímulos sobre um único fundo de tela (sem textura ou pictórico ou fotografia) e um mesmo tempo de exposição (50 ms ou 100 ms ou 150 ms ou 200

ms), dependendo do grupo para o qual o participante foi designado. Depois uma tela em branco era inserida durante 500 ms para impedir a observação de pós-imagem dos estímulos. Em seguida, apresentava-se uma tela para a coleta da resposta. Após a resposta ser dada, uma tela aparecia e a próxima tentativa iniciava-se somente após o participante apertar a tecla de espaço. Isso permitia que o participante tivesse o controle sobre a apresentação dos estímulos.

O participante foi instruído a indicar pelo teclado do computador em que local da tela (na parte superior ou inferior) foi apresentado o maior círculo. Ele foi instruído a responder o mais acuradamente possível e nenhum tipo de *feedback* sobre as respostas foi fornecido. Antes de iniciar o experimento, o participante realizou 10 tentativas como treino para se familiarizar com a tarefa.

Os estímulos foram apresentados através do método psicofísico das escadas duplas (Cornsweet, 1962). Houve, portanto, uma série ascendente em que o tamanho do estímulo teste no início era menor do que o do estímulo padrão e uma série descendente na qual o tamanho do estímulo teste inicial era maior do que o do estímulo padrão. Para evitar que o participante previsse a qual série os estímulos pertenciam, ambas foram intercaladas aleatoriamente.

Foram realizadas duas sessões experimentais e em cada uma foi calculado um ponto distinto da função psicométrica: o ponto de 50 % ou Ponto de Igualdade subjetiva (PIS) através da regra *1up-1down* e o ponto de 70,7% com a regra *1up-2down*. Cada uma das sessões, cuja ordem foi aleatorizada entre os participantes, terminou após 24 reversões. Para atingir este critério eram necessárias em média 100 tentativas, o que demorava 10 minutos.

Análise de dados

A média do tamanho do estímulo teste nas últimas 18 reversões nas sessões *1up-1down* e *1up-2down* foi utilizada para o cálculo do PIS e do ponto 70,7%, respectivamente.

Com esses dois pontos da função psicométrica calculou-se a inclinação desta curva, através do ajuste destes à função cumulativa normal. Os PIS e as inclinações das funções psicométricas foram submetidos a uma análise de variância (ANOVA) entre participantes com duas variáveis independentes [3 fundos de tela (sem textura, gradiente de perspectiva pictórico e gradiente com elementos naturais fotografado) x 4 tempos de exposição (50 ms, 100 ms, 150 ms e 200 ms)].

Resultados

O PIS é uma medida de distorção visual e refere-se ao valor do estímulo teste que na média é estimado como igual ao estímulo padrão. O estímulo teste sempre foi apresentado no campo visual superior, assim PIS com valores abaixo de 100 pixels (estímulo padrão) revelam uma superestimação do estímulo teste e valores acima deste valor indicam subestimação. Este último caso ocorreu em apenas um dos doze grupos experimentais (fotografia-100 ms) e, ainda sim, de maneira sutil (PIS=100,62).

A ANOVA revelou efeitos significativos do fator “fundo” [$F(2,84)=4,355$; $p=0,016$] e para a interação entre os fatores “tempo” x “fundo” [$F(6,84)=3,782$; $p=0,002$], entretanto o fator “tempo” não foi significativo [$F(3,84)=1,681$; $p=0,177$]. Para verificar a natureza desta interação entre “tempo” e “fundo” realizou-se três ANOVAs com o fator “tempo” como variável independente. Esta análise foi realizada com o método recomendado por Howell (2002), no qual o erro quadrado médio da primeira análise é utilizado para o cálculo dos efeitos simples. Para a condição controle (sem gradiente de textura) não foi encontrada nenhuma diferença entre os PIS em função do tempo de exposição dos estímulos [$F(3,28)=0,1865$; $p=0,905$], como pode ser observado na Figura 2.

- Inserir Figura 2 -

Para a condição com o gradiente de perspectiva pictórico foi observada uma diferença significativa [$F(3,28)=3,4053$; $p=0,021$] entre os PIS de acordo com os tempos investigados. O pós-teste de Bonferroni ($p=0,015$) revelou uma grande superestimação do tamanho aos 100 ms ($X=94,59$; $EPM=0,82$) em comparação aos julgamentos realizados aos 200 ms ($X=99,25$; $EPM=0,85$). Entre o tempo de 150 ms ($X=95,56$; $EPM=1,14$) e o de 200 ms foi revelado um efeito marginal ($p=0,083$). Observa-se na Figura 3, que aos 50 ms o valor do PIS está próximo ao de 100 pixels, indicando pouca distorção de tamanho. Por outro lado, a superestimação observada aos 100 ms revela que a percepção de tamanho foi fortemente afetada pela informação de distância presente no gradiente em questão, indicando que ocorreu a integração tamanho-distância. Nos tempos subsequentes, as distorções diminuem e os resultados vão em direção à constância do tamanho.

- Inserir Figura 3 -

Nos julgamentos de tamanho realizados no gradiente de textura fotografado (Figura 4), mais uma vez os tempos de apresentação modularam diferencialmente os tamanhos percebidos [$F(3,28)=5,6532$; $p=0,001$]. O pós-teste de Bonferroni ($p=0,007$) indicou uma diferença entre as comparações de tamanho no tempo de 100 ms ($X=100,63$; $EPM=0,80$) e no de 150 ms ($X=94,30$; $EPM=1,37$), porém também foi encontrado um efeito marginal ($p=0,084$) entre 100 ms e 200 ms ($X=96,05$; $EPM=1,60$). Os resultados indicam pouca distorção de tamanho nas comparações realizadas nos tempos de 50 ms e 100 ms. Aos 150 ms, observou-se superestimação de tamanho, mostrando que neste tempo o gradiente de textura fotografado foi processado e afetou fortemente o tamanho percebido. No tempo de 200 ms, os desvios voltam a diminuir, sugerindo que o PIS em tempos maiores alcançaria valores próximos aos 100 pixels.

- Inserir Figura 4 -

Outro parâmetro importante para analisar os resultados obtidos em tarefas de comparar tamanhos é a inclinação da função psicométrica, a qual oferece informações sobre a sensibilidade do participante em discriminar diferenças entre os estímulos. A ANOVA mostrou efeitos significativos para o fator “tempo” [$F(3,84)=5,641$; $p=0,001$], porém o fator “fundo” [$F(2,84)=0,144$; $p=0,866$] e a interação entre os fatores não foram significativos [$F(6,84)=0,633$; $p=0,704$]. O pós-teste de Bonferroni ($p=0,001$) aplicado aos tempos de exposição revelou que os participantes tiveram uma maior sensibilidade para discriminar a diferença de tamanho entre os estímulos aos 200 ms em comparação aos julgamentos realizados aos 50 ms (Figura 5).

- Inserir Figura 5 -

De maneira geral, nota-se pelas inclinações médias das funções psicométricas, que os participantes discriminaram melhor o tamanho dos estímulos à medida que o tempo de exposição aumentava. Este efeito ocorreu independente do gradiente de textura, indicando que estes não afetaram de maneira diferente a sensibilidade do sistema visual em uma tarefa de comparar tamanhos em apresentações breves.

Discussão

O presente estudo teve por objetivo verificar o tempo requerido para integrar informações de tamanho e distância. Para tal, utilizou-se uma tarefa de comparação de tamanho de objetos inseridos em gradientes de textura pictóricos e fotografados. Os estímulos eram apresentados em tempos muito breves e as distorções de tamanho observadas são indicadoras do processamento subjacente. Assim, uma maior distorção em determinado tempo indicaria a influência dos gradientes de textura, mostrando que ocorreu a integração tamanho-distância.

Os resultados mostraram que o tempo de exposição tem um efeito significativo sobre o tamanho relativo, em função da maneira como o espaço foi representado. Na condição pictórica (perspectiva), foram observadas grandes distorções de tamanho aos 100 ms. Resultados semelhantes foram encontrados por Torro-Alves, Bacci, Bernardino e Fukusima (no prelo). Estes autores apresentavam brevemente (100 ms) duas barras verticais sobre gradientes de textura e linha do horizonte pictóricos. Os dados indicaram uma grande superestimação do tamanho da barra situada no campo visual superior em relação à outra no campo visual inferior. Este resultado evidencia a rapidez do sistema visual para integrar as informações de tamanho e de distância. Esta conclusão é adequada para o gradiente de perspectiva pictórico, porém pode não ser verdadeira para um gradiente de textura fotografado.

Representações pictóricas e fotografadas são processadas de modo semelhante, segundo Watanabe (2006). No entanto, observou-se que para as comparações de tamanho realizadas sob o gradiente de textura fotografado, o processamento é mais lento e há a necessidade de um tempo maior (150 ms) para acontecer a integração tamanho-distância. Este resultado pode ser entendido pela maior complexidade informativa, pois a textura fotografada era composta por elementos naturais, necessitando assim um maior tempo de exposição para ser processada.

Por outro lado, para as comparações de tamanho realizadas na condição controle (sem textura) observou-se uma pequena superestimação em todos os tempos de exposição investigados. Nesta condição, pode-se afirmar que a altura no campo visual foi utilizada como uma informação de profundidade, afetando assim as comparações de tamanho (Epstein, 1966).

Ademais, as distorções de tamanho observadas aos 50 ms são similares tanto na ausência quanto na presença dos gradientes de textura, sejam estes pictóricos ou fotografados. Este resultado indica que neste tempo os gradientes não forneceram informação sobre distância, já que estes não são um indício de profundidade *per se* (Tsutsui et al., 2002). Deste modo, pode-se afirmar que 50 ms não é um tempo suficiente para processar ou mesmo detectar os gradientes de textura em questão.

No que diz respeito à inclinação da função psicométrica, esperava-se encontrar uma relação direta entre sensibilidade para discriminar tamanhos e os tempos de exposição. Isto é, com um tempo maior para inspecionar a imagem, mais recursos cognitivos (*top-down*) são recrutados, o que repercutiria no desempenho da tarefa. Em conformidade com as expectativas, os resultados indicaram que os participantes tiveram uma maior capacidade de discriminação entre os estímulos nas comparações realizadas aos 200 ms em comparação às efetuadas aos 50 ms. Este efeito significativo do tempo de exposição ocorreu independente do fundo de tela.

A capacidade do sistema visual humano em extrair informações de profundidade a partir do gradiente de textura, especialmente da perspectiva, já foi descrita anteriormente (Allison & Howard, 2000; Goodenough & Gilliam, 1997; Wu, He & Ooi, 2007), inclusive em estudos com abordagem neurofisiológica (Di Lollo et al., 2000; Tong 2003). Entretanto, há poucas evidências de que esta rapidez e eficiência para processar os gradientes de textura reflitam na modulação do tamanho percebido, como foi mostrado no experimento aqui descrito.

Nesta direção, pode-se citar o estudo de Champion e Warren (2008). Estes autores com uma tarefa de busca visual investigaram a hipótese de que o tamanho retiniano é acessado diretamente em tarefas que requerem um processamento visual muito rápido. Os

resultados não confirmaram esta hipótese, indicando que o escalonamento do tamanho pode ser realizado em tempos inferiores a 100 ms. Também consonante com estes achados, Farran, Whitaker e Patel (2009) encontraram que é difícil ignorar a adição de dicas de profundidade em uma tarefa de julgamento de tamanhos, sugerindo que a apreensão da informação de distância não necessita de atenção focada.

Em conjunto, estes resultados dão suporte às teorias *bottom-up* de processamento visual, corroborando a abordagem de Gibson (1950), para quem o gradiente de textura fornece informação suficiente e direta para a percepção de tamanho relativo. Desse modo, não é necessário nenhum mecanismo cognitivo ou inferencial entre tamanho e distância, como proposto pela Hipótese da Invariância Tamanho-Distância (Kilpatrick & Ittelson, 1953). Em um estudo recente, Tozawa (2010) também encontrou que o modelo de percepção direta foi mais adequado que a HITD para predizer o tamanho relativo.

Pode-se ainda ressaltar que os resultados aqui descritos mostraram uma diminuição nos desvios das respostas depois dos 100 ms para o gradiente de textura pictórico e após os 150 ms para o gradiente fotografado. Uma hipótese para explicar estes dados é a aplicação de alguma correção cognitiva (por exemplo, alocação da atenção) em tempos maiores. Sabe-se que após 100 ms aumenta a probabilidade de ocorrência de movimentos oculares (Bourne, 2006). Por isso, estudos futuros devem efetuar o controle destes movimentos, especialmente nas condições em que o tempo de exposição dos estímulos é superior a 100ms.

Além disso, há evidências de que os movimentos oculares compartilham vias neurais com a atenção (Awh et al., 2006). Assim, a atenção melhoraria o desempenho dos participantes nesta tarefa, o que é evidenciado pelas menores distorções de tamanho observados em maiores tempos de exposição. No entanto, ainda falta estabelecer a exata influência da atenção sobre a percepção de tamanho, pois alguns trabalhos concluem que a

atenção reduz o tamanho percebido (Tsal, Shalev & Zakay, 2005), enquanto outros afirmam o contrário (Masin 2003). Também é necessário investigar como diferentes processos atentos, tais como a atenção exógena e endógena, afetam tarefas envolvendo percepção visual de tamanho.

Em resumo, encontrou-se que a integração tamanho-distância ocorre em tempos curtos, indicando o envolvimento de mecanismos *bottom-up* no processamento destas informações. Além disso, a integração foi modulada diferencialmente pelo tempo de exposição dos estímulos, de acordo com o gradiente de textura. Na condição pictórica, este processamento tamanho-distância ocorreu em um tempo inferior (100 ms) em relação à observada no fundo fotografado (150 ms).

Referências

- Aks, D. J., & Enns, J. T. (1996). Visual search for size is influenced by a background texture gradient. *Journal of Experimental Psychology: Human Perception and Performance*, 22(6), 1467-1481.
- Allison, R. S., & Howard, I. P. (2000) Temporal dependencies in resolving monocular and binocular cue conflict in slant perception. *Vision Research*, 40 (4), 1869-1885.
- Awh, E., Armstrong, K. M., & Moore, T. (2006). Visual and oculomotor selection: links, causes and implications for spatial attention. *Trends in Cognitive Science*, 10, 124-130.
- Bengston, J. K., Stergios, J. C., Ward, J. L., & Jester, R. E. (1980). Optic array determinants of apparent distance and size in pictures. *Journal of Experimental Psychology: Human, perception and performance*, 6(4), 751-759.

- Bourne, V. J. (2006). The divided visual field paradigm: methodological considerations. *Laterality*, 11, 373-393.
- Champion, R. A., & Warren, P. A. (2008). Rapid size scaling in visual search. *Visual Research*, 48, 1820-1830.
- Cornsweet, T. (1962). The staircase method in psychophysics. *American Journal of Psychology*, 75, 485-491.
- Di Lollo, V., Enns, J. T., Resink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, 129, 481-507.
- Epstein, W. (1966). Perceived depth as a function of relative height under three background conditions. *Journal of Experimental Psychology*, 72, 335-338.
- Farran, E. K., Whitaker, A., & Patel, N. (2009). The effect of pictorial depth information on projected size judgments. *Attention, Perception and Psychophysics*, 71(1), 207-214.
- Foley, J. M., Ribeiro-Filho, N. P., & Da Silva, J. A. (2004). Visual perception of extent and the geometry of visual space. *Vision Research*, 44 (2), 147-156.
- Found, A., & Müller, H. J. (2001). Efficient search for size targets on a background texture gradient: Is detection guided by discontinuities in the retinal-size gradient of items? *Perception*, 30(1), 21-48.
- Gibson, J. J. (1950). *The perception of the visual world*. New York: Houghton Mifflin.
- Gogel, W. C., & Da Silva, J. A. (1987). A two-process theory of the response to size and distance. *Perception & Psychophysics*, 41(3), 220-238.
- Goodenough, B., & Gilliam, B. (1997). Gradients as visual primitives. *Journal of Experimental Psychology: Human Perception and Performance*, 23(2), 370-387.

- Haber R.N., & Levin C. A. (2001). The independence of size perception and distance perception. *Perception & Psychophysics*, 63(7), 1140-1152.
- Hecht, H., van Doorn, A., & Koenderink, J. J. (1999). Compression of visual space in natural scenes and in their photographic counterparts. *Perception & Psychophysics*, 61(7), 1269-1286.
- Howell, D. C. (2002). *Statistical Methods for Psychology*. Boston: PWS-Kent.
- Kilpatrick, F. P., & Ittelson, W. H. (1953). The size-distance invariance hypothesis. *Psychological Review*, 60, 223-231.
- Masin, S. C. (2003). Test of the effect of attention on judged length of a line. *Perceptual and Motor Skills*, 96, 599-606.
- O'Brien, J. M. D. (1997). *Texture and motion cues in slant perception*. Tese de Doutorado, Universidade de Londres, Londres, Reino Unido.
- Murray, S.O., Boyaci, H. & Kersten, D. (2006) The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience*, 9, 429–434.
- Tong, F. (2003). Primary visual cortex and visual awareness. *Nature Review: Neuroscience*, 4, 219-229.
- Torro-Alves, N., Bacci, A. V. F., Bernardino, L. G., & Fukusima, S. S. (no prelo). Efeitos da perspectiva na amplificação do tamanho de objetos que interceptam a linha do horizonte. *Psicologia: Reflexão e Crítica*.
- Torro-Alves, N., & Fukusima, S. S. (2007). Gradientes de textura com linhas de perspectiva e horizonte amplificam a superestimação de tamanho relativo em espaços pictóricos. *Psicologia: Reflexão e Crítica*, 20 (1), 35-42.
- Tozawa, J. (2010). Role of texture gradient in the perception of relative size. *Perception*, 39 (5), 641-660.

- Tozawa, J., & Oyama, T. (2006). Effects of motion parallax and perspective cues on perceived size and distance. *Perception*, 35(8), 1007 – 1023.
- Tsal, Y., Shalev, L., & Zakay, D. (2005). The lengthening effect revisited: a reply to Prinzmetal and Wilson (1997) and Masin (1999). *Psychonomic Bulletin & Review*, 12 (1), 185-190
- Tsutsui, K., Jiang, M., Yara, K., Sakata, H., & Taira, M. (2001). Integration of perspective and disparity cues in surface-orientation-selective neurons of área CIP. *Journal of Neurophysiology*, 86, 2856-2867.
- Watanabe, T. (2006). Geometrical structures of photographic and stereoscopic spaces. *Spanish Journal of Psychology*, 9(2), 263-272.
- Williams, P. A., & Enns, J. T. (1996). Pictorial depth and framing have independent effects on the horizontal-vertical illusion. *Perception*, 25(8), 921- 926.
- Wu, B., He, Z. J., & Ooi, T. L. (2007). The linear perspective information in ground surface representation and distance judgment. *Perception & Psychophysics*, 69(5), 654-672.

Figura 1

Exemplos dos Fundos de Tela

Nota: Condição sem textura (a), condição com gradiente de textura pictórico (b), condição com gradiente de textura fotografado (c). Em todos os exemplos, ambos os círculos têm o mesmo tamanho (100 pixels no tamanho original).

Figura 2

Pontos de Igualdade Subjetiva das Comparações de Tamanho na Condição Controle (Sem Textura)

Nota: Pontos de Igualdade Subjetiva (PIS) médios para comparações de tamanho realizadas na condição controle nos 4 tempos de exposição investigados (50 ms, 100 ms, 150 ms e 200 ms). As barras de erro referem-se ao erro padrão da média. Não houve diferença significativa entre os tempos de apresentação dos estímulos ($p=0,905$).

Figura 3

Pontos de Igualdade Subjetiva das Comparações de Tamanho no Gradiente de Textura Pictórico

Nota: Pontos de Igualdade Subjetiva (PIS) médios para comparações de tamanho realizadas na situação com o gradiente de textura pictórico nos 4 tempos de exposição investigados (50 ms, 100 ms, 150 ms e 200 ms). As barras de erro referem-se ao erro padrão da média. Houve diferença significativa entre os tempos de exposição ($p=0,021$).

* O PIS foi significativamente menor aos 100 ms do que aos 200ms (pós-teste de Bonferroni, $p=0,015$)

Figura 4

Pontos de Igualdade Subjetiva das Comparações de Tamanho no Gradiente de Textura
Fotografia

Nota: Pontos de Igualdade Subjetiva (PIS) médios para comparações de tamanho realizadas na condição gradiente de textura fotografado nos 4 tempos de exposição investigados (50 ms, 100 ms, 150 ms e 200 ms). As barras de erro referem-se ao erro padrão da média. Houve diferença significativa entre os tempos de exposição ($p=0,001$).

* O PIS foi significativamente menor aos 150 ms do que aos 100ms (pós-teste de Bonferroni, $p=0,007$)

Figura 5

Inclinação da Função Psicométrica das Comparações de Tamanho nos Doze Grupos Experimentais

Nota: Inclinação média das funções psicométricas para comparações de tamanho realizadas nas três condições de fundo de tela (controle, gradiente de textura pictórico e gradiente de textura fotografado) e nos 4 tempos de exposição investigados (50 ms, 100 ms, 150 ms e 200 ms). As barras de erro referem-se ao erro padrão da média. Houve diferença significativa entre os tempos de apresentação dos estímulos ($p=0,001$).

* A Inclinação foi significativamente menor aos 50 ms do que aos 200ms (pós-teste de Bonferroni, $p=0,001$)

ANEXO A - Termo de Consentimento Livre e Esclarecido

Gostaria de convidá-lo a participar da pesquisa que estou desenvolvendo para o meu trabalho de doutorado no programa de Psicobiologia da FFCLRP. Esta pesquisa tem por objetivo investigar como informações de profundidade afetam o desempenho em se comparar visualmente o tamanho de círculos.

O requisito para que você participe desta pesquisa é ter boa acuidade visual. Por isso, você já fez um teste de acuidade visual, a qual se constatou ser normal em ambos os olhos.

Caso você queira participar como voluntário desta pesquisa sua tarefa será a seguinte: indicar dentre dois círculos, qual é o maior. Em uma sala escura, os círculos serão apresentados num monitor de computador e suas respostas serão coletadas através do teclado do mesmo.

O experimento não apresenta riscos à integridade de sua saúde física ou mental e terá uma duração de aproximadamente 10 minutos.

É importante que você finalize o experimento, porém, se desejar interrompê-lo ou encerrá-lo, você assim poderá proceder a qualquer momento, sem qualquer prejuízo a você. Garanto esclarecer suas dúvidas sobre a pesquisa em qualquer aspecto e momento que desejar.

Os dados desta pesquisa serão divulgados em reuniões e publicações científicas e as identidades dos participantes serão mantidas em absoluto sigilo.

Este termo foi elaborado em duas vias, sem do que uma das cópias será fornecida a você e a outra será arquivada por mim.

Declaro que estou ciente das informações acima e concordo em participar da pesquisa.

Local e data: _____

Nome do participante: _____

Endereço: _____

Assinatura do participante: _____

Nome do Pesquisador: Leonardo Gomes Bernardino

Assinatura do pesquisador: _____

Orientador: Prof. Dr. Sérgio Sheiji Fukusima

Assinatura do orientador: _____

Endereço: Av. Bandeirantes, 3900. Bairro: Monte Alegre. Cep: 14.040-901. Ribeirão Preto-SP
Universidade de São Paulo - Faculdade de Filosofia, Ciências e Letras.
Departamento de Psicologia e Educação. Telefone: (16) 3602-3729/3602-4448

ANEXO B - Parecer do Comitê de Ética em Pesquisa (CEP) da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto de Ribeirão Preto da Universidade de São Paulo.

**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
COMITÊ DE ÉTICA EM PESQUISA - CEP**

Of.CEtP/FFCLRP-USP. 089/2009-27/11/2009

Senhor Pesquisador:

Comunicamos a V. Sa. que o trabalho intitulado "Tamanhos pictóricos percebidos sobre gradientes de textura desenhados e fotografados após exposições breves", foi re-analisado pelo Comitê de Ética em Pesquisa da FFCLRP-USP, e, enquadrado na categoria: **APROVADO**, de acordo com o Processo CEP-FFCLRP nº 456/2009 – 2009.1.1558.59.5.

Atenciosamente,

Profa. Dra. ANA RAQUEL LUCATO CIANFLONE
Coordenadora do Comitê de Ética em Pesquisa – FFCLRP-USP

Ilustríssimo Senhor
LEONARDO GOMES BERNANRDINO
Aluno(a) do Programa de Pós-Graduação em Psicologia
Desta FFCLRP-USP

c.c.: Prof(a). Dr(a).
Prof. SERGIO SEIJI FUKUSIMA
DOCENTE DO DPE-FFCLRP

CEP-FFCLRP-USP - Fone: (016) 602-3653 - Fax: (016) 633-5015
Avenida Bandeirantes, 3900 - Bloco A - 14040-901 - Ribeirão Preto - SP - Brasil

ANEXO C - Instruções aos participantes – Experimento 1

O experimento inicia-se com a seguinte instrução que aparecerá na tela do computador:

Agora será apresentada uma série de pares de círculos no eixo vertical e você deverá indicar qual é o maior.

Se o círculo acima na tela for o maior, aperte a tecla M. Se o maior for o da parte inferior, aperte a tecla Z.

O seu queixo deve permanecer apoiado durante todo o experimento.

O seu olhar deve sempre se manter fixo no centro da tela.

A seguir, dez exemplos da tarefa.

Nesse momento, serão apresentados dez pares de estímulos para o participante para que ele conheça a tarefa antes do início da coleta das respostas. Então aparecerá a seguinte tela:

Quando estiver pronto pra começar é só apertar a tecla P.

Ao final da primeira sessão, uma tela será apresentada:

Nesse momento, você completou metade do experimento.

Se quiser sair da sala, descansar, beber água, tudo bem.

Chame o experimentador para o reinício do experimento.

O participante é quem resolve o momento do reinício do experimento, isto é, o tempo entre a primeira e a segunda sessão é por ele definida. No reinício, a tela a seguir é apresentada:

Lembre-se de manter o olhar fixo no centro da tela e apoiar o queixo.

Quando estiver pronto, aperte qualquer tecla para reiniciar o experimento.

Ao final do experimento, uma tela de agradecimento é mostrada:

Muito obrigado, sua participação foi muito importante!

ANEXO D - Normas do Comitê de Ética da Universidade de Barcelona**COMITES DE BIOÉTICA DE LAS UNIVERSIDADES PUBLICAS**
NORMAS DE FUNCIONAMIENTO**EXPOSICION DE MOTIVOS**

La normativa contenida en diferentes convocatorias de financiación de la investigación exige que los proyectos que impliquen investigación con seres humanos, utilización de datos personales, muestras biológicas de origen humano, experimentación animal, agentes biológicos o empleen organismos genéticamente modificados, no sólo cumplan los requisitos establecidos en cada caso por la legislación sino que también cuenten con la autorización expresa emitida por el Comité de Ética del Centro en que se vaya a realizar la investigación.

La Universidad de Barcelona, como Organismo Público de Investigación, es consciente de su deber y responsabilidad ante la sociedad y, en consecuencia, constituye un Comité de Bioética que de forma ágil y efectiva proporcione respuesta a las necesidades actuales así como a nuevos problemas que puedan surgir. Sus funciones composición y normas básicas de funcionamiento son las siguientes:

1. FUNCIONES DE LOS COMITÉS DE BIOÉTICA DE LA UNIVERSIDAD

1. Elaborar informes solicitados por organismos e investigadores sobre proyectos o trabajos de investigación cuando se aborden estudios con seres humanos, se utilicen muestras biológicas de origen humano, experimentación animal, agentes biológicos y organismos genéticamente modificados.
2. Valorar cualquier otro proyecto de investigación que pueda afectar de modo directo los derechos fundamentales de las personas e intereses vinculados a la defensa y la protección del medio ambiente.
3. Emitir informes sobre problemas y conflictos éticos que puedan suscitar la investigación y la docencia.
4. Velar, en la medida de lo posible, por el cumplimiento de las buenas prácticas de investigación y experimentación.
5. Difundir y promover debates en la comunidad universitaria sobre cuestiones bioéticas de interés general.
6. Empezar iniciativas de divulgación en bioética, promoviendo la información precisa para fomentar la comprensión y el diálogo en el ámbito público.

2. COMPOSICIÓN

El Comité de Bioética estará integrada, al menos, por los miembros o asesores siguientes:

- Un especialista en ética médica.
- Un investigador en ciencias de la vida, de la salud o del medio ambiente.
- Un antropólogo, sociólogo, historiador o psicólogo, investigador en ciencias humanas o sociales.
- Un jurista.
- Un especialista en ética o filosofía política.
- Un especialista en bioética de las escuelas universitarias de enfermería o trabajo social.

El Comité estará presidido por el Rector (o persona en quien delegue), el cual designará un secretario entre sus miembros.

3. DURACIÓN

La composición del Comité tendrá una duración de tres años y será renovable por mitades.

4. NORMAS GENERALES DE FUNCIONAMIENTO

1. El funcionamiento de este Comité estará sometido a lo previsto en el presente texto y en todo lo no especificado aquí se estará a lo regulado en la Ley 30/92 de Régimen Jurídico de las Administraciones Públicas y del Procedimiento administrativo común.
2. El Comité actuará siempre que lo requieran sus funciones y a solicitud del Presidente o Secretario.
3. El Comité trabajará en Pleno y en subcomisiones. Las subcomisiones estarán formadas como mínimo por dos miembros y realizarán cuantas funciones les sean delegadas por el pleno. El Pleno podrá recabar para sí el conocimiento de cualquier asunto.
4. Para la adopción válida de decisiones se requerirá la asistencia mínima de la mitad mas uno de los miembros. Las decisiones se adoptarán por mayoría.
5. Los miembros del Comité estarán plenamente obligados por el principio de confidencialidad, tanto de los debates como de los informes antes de que se hagan públicos.
6. Los miembros del Comité no podrán participar en las deliberaciones que afecten a proyectos en que participen como investigadores ni en otros casos en que puedan presentarse conflictos de intereses.
7. Cuando el Comité lo considere oportuno podrá recabar la opinión de expertos externos, que tendrán también que respetar el principio de confidencialidad.
8. El Comité tiene competencia para solicitar a los investigadores de un proyecto objeto de evaluación cuanta información adicional considere necesaria.

UNIVERSITAT DE BARCELONA**COMISSIÓ DE BIOÈTICA**

Declaració de no existència d'experimentació en éssers humans ni la utilització de mostres biològiques d'origen humà.

El/La sotaginant , investigador/a principal del projecte intitulat

DECLARO QUE:

El projecte abans ressenyat no implica cap tipus d'experimentació en éssers humans ni la utilització de mostres biològiques d'origen humà.

El que faig constar a l'objecte que la Comissió de Bioètica de la Universitat de Barcelona , emeti l'informe corresponent.

Nom i Cognoms

Dept.

Adreça Postal

Telèfon

Fax

Correu Electrònic

Signatura:

Barcelona,

ANEXO E - Instruções aos participantes – Experimento 2

Instrucciones del Experimento

Vas a hacer un experimento sobre percepción de tamaño, que tiene una duración de 15 minutos, aproximadamente.

En cada presentación, serán expuestos dos círculos en la pantalla del ordenador, de manera simultánea, uno arriba y otro abajo del *punto de fijación*, que es una cruz que se muestra en el centro de la pantalla.

Tu tarea consistirá en determinar la posición en que fue presentado el círculo más grande.

→ Si tú crees que el círculo más grande fue presentado arriba del punto de fijación, deberás pulsar ↑ en el teclado numérico.

→ Por el contrario, si crees que el círculo más grande fue presentado abajo del punto de fijación, deberás pulsar ↓ en el teclado numérico.

Mientras sólo el punto de fijación esté en la pantalla, deberás mirarlo. Por otro lado, cuando los círculos son presentados, la exploración visual es libre.

Es importante decir que la presentación de los círculos es muy breve. Una cámara controlará dónde fijas la mirada en la pantalla.

*

*

*