Proceedings of the
38th Annual Meeting of the International Society for Psychophysics

FECHNER DAY 2022

Lund, Sweden.
23 – 26 August.

Reference
G. R. Patching (Ed.), *Fechner Day 2022 – Proceedings of the 38th Annual Meeting of the International Society for Psychophysics*. Lund, Sweden: International Society for Psychophysics.

Acknowledgements
Financial administration: Cristina Svensson.
Logistics: Anna Håkansson Öreberg.

Distribution
International Society for Psychophysics
C/O Geoffrey R. Patching
Department of Psychology
Lund University
SE-221 00, Lund
Sweden.
E-mail: Geoffrey.Patching@psy.lu.se

International Society for Psychophysics

Founded in Cassis, France, 22 October 1985

Executive Committee

President: Kazuo Ueda (*Kyushu University, Japan*)
Vice President: Leah Fostick (*Ariel University, Israel*)
Secretary: Jordan Schoenherr (*Carleton University, Canada*)
Treasurer: Wolfgang Ellermeier (*Technical University of Darmstadt, Germany*)

Non-Executive Scientific Communications Officer:
Natalia Postnova (*Kyushu University, Japan*)

Program and Organization Committee, Fechner Day 2022

Geoffrey R. Patching (*Lund University, Sweden*)
Anette Lidström (*Lund University, Sweden*)
Andrey Nikolaev (*Lund University, Sweden*)
Maarit Johnson (*Stockholm, Sweden*)
Åke Hellström (*Stockholm University, Sweden*)

Past Annual Meetings

Year	Venue	Organizer(s)
1985	Marseille, France	G. Canévet and B. Scharf
1986	Cassis, France	A. M. Bonnel, G. Canévet, C.-A. Possamaï, B. Scharf, B. Berglund, U. Berglund, and R. Teghtsoonian
1987	Durham, NH, USA	M. Teghtsoonian and R. Teghtsoonian
1988	Stirling, Scotland	H. Ross, R. Macdonald, C.-A. Possamaï, R. Teghtsoonian, M. Treisman, and R. Warren
1989	Cassis, France	A. M. Bonnel, G. Canévet, C.-A. Possamaï, and B. Scharf
1990	Würzburg, Germany	F. Müller
1991	Durham, NH, USA	G. R. Lockhead
1992	Stockholm, Sweden	G. Borg, Å. Hellström, and G. Neely
1993	Mallorca, Spain	A. Garriga-Trillo, P. R. Miñón, C. Carcía-Gallego, P. Lubin, J. M. Merino, M. J. Rubio-Gómez, and A. Villario
1994	Vancouver, Canada	L. M. Ward
1995	Cassis, France	C.-A. Possamaï, H. Ross, B. Scharf, R. Teghtsoonian, and L. M. Ward
1996	Padua, Italy	S. C. Masin
1997	Poznan, Poland	R. Gotebiewski, E. Hojan, T. Hornowski, P. Kokowski, A. Majchrzak, P. Miecznik, M. Labowski, P. Pekala, A. Preis, E. Skrodzka, E. Wichlinska, and A. Wicher
1998	Québec, Canada	S. Grondin, Y. Lacouture, and R. Rousseau
1999	Tempe, AZ, USA	P. R. Killeen and W. R. Uttal
2000	Strasbourg, France	C. Bonnet
2001	Leipzig, Germany	E. Sommerfeld, R. Kompass, and T. Lachmann
2002	Rio de Janeiro, Brazil	J. A. da Silva, E. H. Matsushima, and N. P. Ribero-Filho
2003	Larnaca, Cyprus	B. Berglund and E. Borg
2004	Coimbra, Portugal	A. Olivera, M. Teixeira, G. F. Borges, and M. J. Ferro
2005	Traverse City (MI), USA	J. S. Monahan, S. M. Sheffert, and J. T. Townsend
2006	St Albans, UK	D. E. Kornbrot, R. M. Msetfi, and A. W. McRae
2007	Tokyo, Japan	S. Mori, T. Miyaoka, and W. Wong
2008	Toronto, Canada	B. A. Schneider, B. M. Ben-David, S. Parker, and W. Wong
2009	Galway, Ireland	M. A. Elliott, S. Antonijević, S. Berthaud, P. Mulcahy, C. Martyn, B. Bargary, and H. Schmidt
2010	Padua, Italy	A. Bastianelli, G. Vidotto, L. Filipponi, D. Polezzi, A. Spoto, D. Massidda, and M. Scandola
2011	Herzliya, Israel	D. Algom, D. Zakay, E. Chajut, S. Shaki, Y. Mama, and V. Shakuf

2012	Ottawa, Canada	S. R. Carroll, E. Gallitto, C. Leth-Steensen, W. M. Petrusic, J. R. Schoenherr, and D. A. Verger
2013	Freiburg, Germany	J. Wackermann, M. Wittmann, and W. Skrandies
2014	Lund, Sweden	G. R. Patching, Å Hellström, M. Johnson, E. Borg, and R. Bååth
2015	Québec, Canada	S. Grondin, V. Laflamme, G. Plante, L. Azari, and M.-L. Roy
2016	Moscow, Russia	I. G. Skotnikova, A. N. Gusev, G. Y. Menshikova, I. Blinnikova, O. Korolkova, S. A. Emelianova, V. M. Shendyapin, V. E. Doubrovski, A. I. Kovalev, A. Kremlev, and N. N. Volkova
2017	Fukuoka, Japan	Y. Nakajima, K. Ueda, and G. B. Remijn
2018	Lüneburg, Germany	F.Müller, L. Ludwigs, and M. Kupper
2019	Antalya, Turkey	M. A. Elliott, E. Gülbetekin, S. Bayraktar, E. V. Özsoy, S. Arndt, N. du Bois, A. Civilidag, M. Hekimoglu, B. Demir, and A. Ozen
2020	Ottawa, Canada (online)	J. R. Schoenherr, T. Hubbard, W. Stine, and Leth-Steensen
2021	Ottawa, Canada (online)	J. R. Schoenherr, T. Hubbard, and K. Ueda

Preface

Welcome to Fechner Day 2022 - the first onsite meeting since restrictions were introduced to contain the spread of the COVID-19 virus in 2020. For the last two years, the Fechner Day meeting has been held online via Zoom. There are some benefits of meeting via Zoom, but we hope everyone agrees that the experience of meeting in-person is unrivalled. This year we have a great scientific program with off-camera, uninhibited, time for socialization and stimulating discussion of scientific works. Again, we find older, long standing, members of the ISP supporting the future generation of young researchers. The yearly Fechner Day conference is not just about listening to great presentations, but actively taking part in discussions that encourage inspiration, new ideas, and innovative thinking. The social interaction afforded by an onsite conference provides a welcome forum for the advancement of the science of psychophysics. It is an honor to host this year's 38th Annual Meeting of the International Society for Psychophysics. A warm in-person welcome to the beautiful city of Lund in the province of Skåne, southern Sweden.

/Geoff Patching

ABSTRACTS AND PAPERS

:: Listed Alphabetically by First Author ::

THE ROLE OF FEEDFORWARD AND RECURRENT PROCESSING DURING ULTRA-RAPID AND MAN-MADE OBJECT IMAGE CATEGORIZATION: A DISTRIBUTIONAL REPONSE TIME STUDY

Ann-Kathrin Beck[1], Joana Costa do Carmo[2], and Thomas Lachmann[1]
[1] *Center for Cognitive Science, University of Kaiserslautern, Kaiserslautern, Germany*
[2] *Centro de Investigação em Ciências Psicológicas, Faculdade de Psicologia, Universidade de Lisboa, Portugal*
<*Ann-Kathrin.Beck@sowi.uni-kl.de*>

Abstract

This study examined the effects of natural vs. man-made object images, presentation duration, and typicality level on the shapes of reaction time and accuracy distributions in an ultra-rapid superordinate categorization paradigm. In contrast to mean performance measures, the distributional analyses show that (a) in fast responses (emitted around 300 ms), which are presumably based on feedforward-only processing, the advantage for natural over man-made object images is present (which is consistent with previous reports), that (b) in slower responses (around 450 ms) this effect reverses, demonstrating an advantage for man-made object over natural images. Our results emphasize the idea that feedforward-only processing are sufficient to perform an ultra-rapid superordinate categorization and that natural images benefit more from these feedforward-only processing than man-made object images.

THE EFFECT OF THE EMOTIONAL CONTENT OF IMAGES DURING ENCODING AND REPRODUCTION PHASES IN AN INTERVAL REPRODUCTION TASK: A COMPARISON BETWEEN CONTROL, ASD, AND ADHD GROUPS

Sonya Bernard[1], Pier-Alexandre Rioux[1], Karine Morasse[2], and Simon Grondin[1]
[1]*Université Laval, Québec, Canada*
[2]*Centre intégré de santé et de services sociaux de Chaudière-Appalaches, Lévis, Canada*

Abstract

This study addresses the impact of the presentation of faces expressing anger on the capability of teenagers with ASD or ADHD to reproduce 2- and 10-sec time intervals. Either a neutral or an angry face was presented during the encoding and reproduction phases of the experiment. At 2 sec, there is an over-reproduction of the duration when an angry face is presented during the encoding phase, this effect being the same for the three groups (ASD, ADHD, and control). At 10 sec, the presentation of an angry face during the encoding phase produces an effect only when a neutral face is presented during reproduction. With regard to the coefficients of variation (CV), the variability is smaller (smaller CV) in ASD participants when an angry face rather than a neutral face is presented during the reproduction phase, while the variability is smaller in ADHD participants when a neutral face rather than an angry face is presented during the reproduction phase.

This study focuses on the processing of temporal information in people aged 11 to 18 who have been diagnosed either with autism spectrum disorder (ASD) or attention deficit/hyperactivity disorder (ADHD). It is known that these disorders might lead to difficulties in the processing of temporal information (Casassus et al., 2019; Toplak et al., 2006; Walg et al. 2015; see Grondin, 2020). It is also known that when face presentations delimit the duration of an interval to be estimated, the emotion expressed by these faces influences their perceived duration (Gil, & Droit-Volet, 2011; Grondin et al., 2015; Ouellet et al., in press) and that people with ASD or ADHD may experience difficulty in reading the emotional content of facial expressions (Yuill & Lyon, 2007).

To investigate these characteristics, the present study uses an interval reproduction task (Mioni et al., 2014). This type of task includes two phases: one dedicated to the encoding of the duration to be reproduced, and one dedicated to the reproduction. In the present study, the time intervals to be reproduced were delimited by visual stimuli consisting of the face of a person expressing either a neutral or an angry expression. More specifically, we wanted to know if the impact of presenting these stimuli (neutral vs. angry) at the encoding phase would depend on their presentation during the reproduction phase.

The effect of neutral vs. emotional stimuli in an interval reproduction task may vary with the efficiency in reading emotional expressions. This ability can vary among different people. Therefore, an additional aim of this study was to determine whether the effect of emotional facial expressions, if any, when presented during the encoding or reproduction phase, would also manifest itself in people with ASD or ADHD, i.e., in people that may experience some difficulty in reading facial expressions, or who may have a different reaction from that of control participants when facing an expression of anger (e.g. Eack et al., 2015; Tehrani-Doost et al., 2017).

Method

Participants

Eleven persons with ASD (*M* age = 14.1; 9 boys and 2 girls; 7 with comorbid ADHD), 11 persons with ADHD (*M* age = 14.0; 7 boys and 4 girls), and 10 neurotypical persons (*M* age = 14.1; 6 boys and 4 girls) participated in the experiment. No participants had intellectual deficits.

Material

The experiment was mainly conducted at the Laboratory for Research in Psychology of Perception at Université Laval, where tasks were conducted on a Toshiba Satellite A 70-RX1F Pentium IV 538 (3,2 gHz) computer, using version 2.0 of E-Prime version 2. Some tasks were conducted outside the laboratory on a portable computer, Inspiron 1521 de DELL, also using E-Prime version 2.0.

Faces presented during the interval reproduction task were drawn from a bank of videos of standardized dynamic facial expressions of pain and emotions (Simon et al., 2008). A screenshot using Capture (Windows Vista) was taken of the last image of the videos, where anger is expressed, to produce the static images used in the study.

Participants also completed a series of tests to assess their cognitive capabilities and their level of social functioning. Cognitive capabilities were assessed with the *Conners' Continuous Performance Test* (CPT-II; Conners, 2000); the WISC-IV was used for participants below 17, the WAIS-IV was used for participants aged 17 and 18, and the WASI was used in the control group. Social functioning was assessed with the *Social Skills Improvement System* (SSIS; Gresham & Elliott, 2008), whereas the *Social Communication Questionnaire* (SCQ; Rutter et al., 2003) was used to detect ASD symptoms in the ADHD and control groups (exclusion criterion).

Task

Participants were conducted into different conditions of time perception tasks (including temporal bisection). Only the conditions related to the interval reproduction task are reported here. Intervals to be reproduced lasted 2 or 10 sec. The task consisted in evaluating the duration of the presentation of a face that appeared on the screen in order to then reproduce that duration. Participants would press the spacebar when they believed that the elapsed time matched the time previously presented, causing the face to disappear from the screen. Participants were asked to pay attention to the face on the screen during the encoding and reproduction phases and to avoid explicit counting.

Participants were invited to read the instructions in the presence of the experimenter and to ask any questions if necessary. Four practice trials were then carried out in the presence of the experimenter, after which the experimental trials began. The experimental session was divided into two blocks, one for each target duration (2 and 10 sec), and the order of presentation of the target durations was counterbalanced between participants. Each block was subdivided into four conditions of 10 trials, with two congruent conditions and two incongruent conditions. The congruent conditions presented the same emotion in both phases of the same trial. Therefore, an angry (or neutral) face appeared during both the encoding and reproduction phases of the same trial (neutral-neutral and angry-angry). The incongruent conditions presented different emotions within the same trial (neutral-angry and angry-neutral). These conditions were presented randomly. The faces of four actors (2 men and 2 women) were presented during the task. It should be noted that different faces were presented in both phases of the same trial.

Each trial began with the instruction “Evaluate the duration of presentation of the face that follows” in the center of the screen. This instruction was followed by a face for which the participant had to estimate the duration of presentation (encoding phase). The instruction “Reproduce the estimated duration” then appeared on the screen, followed by a new face. Participants then pressed the spacebar when the elapsed time seemed to correspond to that previously presented.

After each trial, the instruction on the screen “Press space bar to continue” allowed participants to take a break, if needed. A 5-min break was also suggested between blocks. Once the task was completed, participants were asked about any strategies used to estimate and reproduce the duration, as well as about the emotions perceived throughout the task.

Results

For each experimental condition and target duration, there are two dependent variables of interest, one associated with accuracy, the other with precision. Accuracy was obtained using the mean reproduced duration minus the target duration. As for precision, it is represented by the coefficient of variation (CV), which is the variability (one standard duration) of the reproductions divided by the average reproduction. The data (reproduced intervals) were first filtered to eliminate outliers. In the ASD, ADHD, and control groups, 5,91%, 4,32%, and 3,5% of data, respectively, were not included in the final analyses. Also, because the postulates of homogeneity of variances and normality of data distribution were violated, the Aligned Rank Transformation procedure was used. Finally, a 3 (group ASD, ADHD, control) × 2 (anger vs. neutral at the encoding phase) × 2 (anger vs. neutral at the reproduction phase) with repeated measures on the last two factors was conducted for each target duration condition, that is, 2 and 10 sec.

Constant error (CE)

At 2 sec, there is a significant Encoding effect, $F(1,29) = 11{,}913$, $p = 0{,}002$, $\eta_p^2 = 0{,}29$ (*Figure 1*). The CE is significantly higher when the face presented during the encoding phase expresses anger. This tendency is observed in the three groups and does not depend on the face presented during the reproduction phase. There is no significant Reproduction effect, $F(1,29) = 1{,}709$, $p = 0{,}201$, $\eta_p^2 = 0{,}056$, and no Group effect, $F(2,29) = 0{,}695$, $p = 0{,}507$, $\eta_p^2 = 0{,}046$. All interactions are not significant (all *p*s > 0,26).

As for the CE at 10 sec, both Encoding, $F(1,29) = 0{,}709$, $p = 0{,}407$, $\eta_p^2 = 0{,}024$, and Reproduction effects, $F(1,29) = 0{,}780$, $p = 0{,}384$, $\eta_p^2 = 0{,}026$, are not significant, but their interaction is, $F(1,29) = 5{,}955$, $p = 0{,}021$, $\eta_p^2 = 0{,}170$ (*Figure 2*). When the face expresses anger during the encoding phase, the CE is higher when the face during the reproduction phase shows a neutral expression rather than anger ($p = 0{,}023$). When the face is neutral during the encoding phase, the face’s expression (neutral vs. anger) during the reproduction phase does not affect the CE. Neither the Group effect, $F(2,29) = 0{,}2432$, $p = 0{,}786$, $\eta_p^2 = 0{,}016$, nor the Interaction effects are significant: Encoding × Group, $F(2,29) = 1{,}174$, $p = 0{,}323$, $\eta_p^2 = 0{,}075$; Reproduction × Group, $F(2,29) = 0{,}011$, $p = 0{,}989$, $\eta_p^2 = 0{,}001$; and Encoding × Reproduction × Group, $F(2,29) = 2{,}896$, $p = 0{,}071$, $\eta_p^2 = 0{,}166$.

Coefficient of variation (CV)

At 2 sec, neither the Main effect (all *p*s > .12) nor the Interaction effects (all *p*s > .26) were significant. At 10 sec, there is no main effect of Encoding, $F(1,29) = 2{,}238$, $p = 0{,}145$, $\eta_p^2 = 0{,}072$, of Reproduction, $F(1,29) = 0{,}0001$, $p = 0{,}991$, $\eta_p^2 = 0{,}000$, nor of Group, $F(2,29) = 2{,}556$, $p = 0{,}095$, $\eta_p^2 = 0{,}150$. There is no Encoding × Reproduction interaction, $F(1,29) = 0{,}292$,

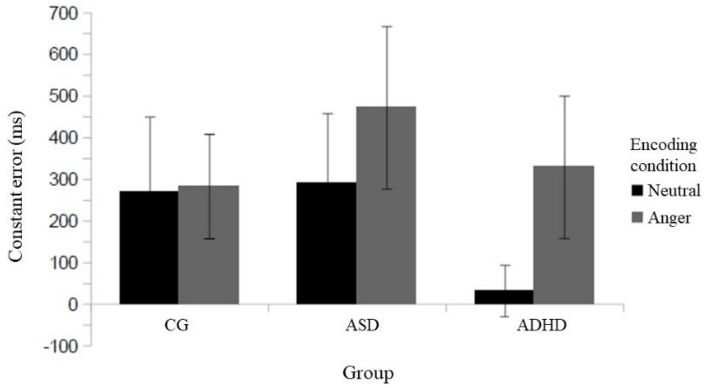


Fig. 1. Constant error (ms) as a function of the encoding condition for each group in the 2000-ms condition. *Note.* Bars are standard error.

$p = 0{,}593$, $\eta_p^2 = 0{,}010$, and no Encoding × Group interaction, $F(2,29) = 0{,}045$, $p = 0{,}956$, $\eta_p^2 = 0{,}003$. However, the Reproduction × Group interaction is significant, $F(2,29) = 3{,}704$, $p = 0{,}037$, $\eta_p^2 = 0{,}203$ (*Figure 3*), and note the Encoding × Reproduction × Group interaction, $F(2,29) = 3{,}154$, $p = 0{,}058$, $\eta_p^2 = 0{,}179$.

For the ADHD group, the CV is significantly higher when the face presented during the reproduction phase expresses anger rather than a neutral expression, and the reversed pattern is observed for the ASD and control groups. Indeed, while the three groups did not differ when presented with neutral faces in the reproduction phase, the CV is much higher for ADHD participants than for participants of the control and ASD groups when anger is presented during the reproduction phase.

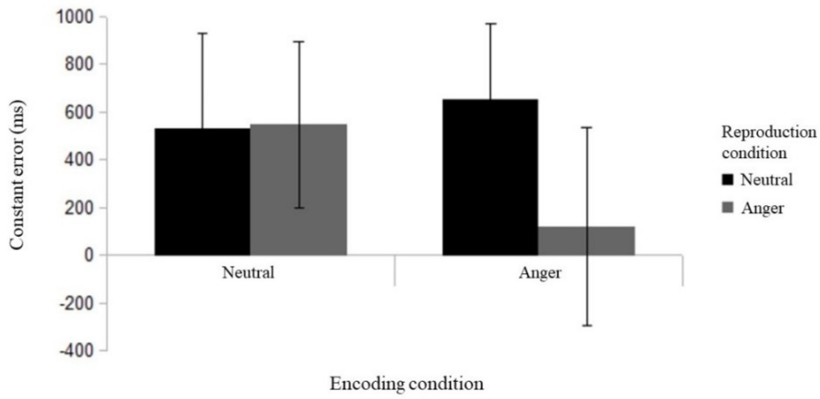


Fig. 2. Mean constant error (ms) for all participants as a function of the encoding and reproduction condition in the 10 000-ms condition. *Note.* Bars are standard error.

Correlational analyses

A series of correlational analyses were conducted to explore the link between the capability to reproduce intervals and some individual characteristics. There is a close relationship between IQ and the CV on the 2-sec reproduction task for ADHD participants (r = -.75), but not for the control (r = -.36) nor ASD participants (r = -.01). For ADHD participants, IQ is also highly correlated (r = -.68) with the CV at 10 sec. In the ASD group, there is a significant link between the working memory index and the CE at 10 sec (r = -.76). We also note the often-close links between omissions and certain indices of temporal performance in ADHD, namely the CE at 2 sec. (r = .61) and the CV at 10 sec. (r = .69).

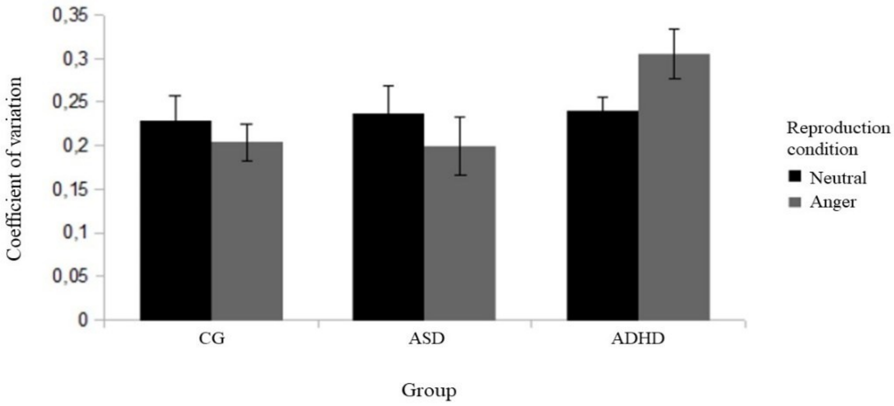


Fig. 3. Coefficient of variation as a function of the reproduction condition for each group in the 10 000-ms condition. *Note.* Bars are standard error.

It is interesting to note that empathy is closely linked with certain indices of temporal performance in control and ADHD participants. The temporal index considered here is the difference between the anger and neutral conditions during the encoding (Encoding A-N) and reproduction (Repro A-N) phases. In the control group, empathy is correlated to the CE at 10 sec in the Repro A-N condition (r = .64). With ADHD participants, empathy is strongly linked to the CE at 2 sec in the Encoding A-N condition (r = -.67), and the CV at 2 sec in the Repro A-N condition (r = .80). We also note a significant correlation in the control group (r = -.68) between self-control and the CV at 2 sec in the Repro A-N condition. Finally, note that in the three groups, the performance of participants who recorded high levels of Hyperactivity/Inattention symptoms tended to be more affected by the presentation of an angry face than those who scored lower on that scale.

Discussion

This study suggests that it is possible to observe the impact of emotion on temporal perception with a reproduction task, even though it requires more cognitive resources than the bisection task that is typically used. For the 2-sec reproduction, the presentation of an angry face during the encoding phase is associated with an over-reproduction of duration in the three groups. An over-reproduction is once again observed when an angry face is presented during the encoding phase in the 10-sec condition, but only when people are exposed to a neutral face during the reproduction phase. The use of static pictures of people expressing anger seems to induce a lower increase of the arousal level during this longer condition. With longer durations, it is possible that people become habituated to the stimulus and/or they succeed in regulating their heightened arousal.

Finally, it appears that inattention tends to be associated with an over-reproduction of duration. Referring to the pacemaker-accumulator mechanism (Internal clock: see Grondin et al., 2015), this over-reproduction may be attributed to an opening of a switch process and a subsequent loss of pulses emitted by the pacemaker, whereas better capacities in working memory were associated with more accurate reproductions. The impact of emotion seems to be modulated by social skills; empathy, which requires cognitive resources that cannot be spent on the timing task; and self-control, which contributes to decreasing arousal. Behavioral problems are rather associated with heightened arousal and ineffective emotion regulation processes, thus leading to a greater temporal distortion with faces expressing anger.

Acknowledgements

This work is part of the Ph. D. thesis of SB under supervision of KM and SG. We would like to thank Vincent Laflamme for his help in data analyses, and Elsa Gilbert, Isabelle Denis, and Renaud Brochard for their comments on this study.

References

Casassus, M., Poliakoff, E., Gowen, E., Poole, D., & Jones, L. A. (2019). Time perception and autistic spectrum condition: A systematic review. *Autism Research. 12*(10), 1440-1462.

Conners, K. (2000). *Conners' Continuous performance test II*. Toronto: Multi Health System.

Eack, S. M., Mazefsky, C. A. & Minshew, N. J. (2015). Misinterpretation of facial expressions of emotion in verbal adults with autism spectrum disorder. *Autism*, *19*, 308-315.

Gil, S., & Droit-Volet, S. (2011). "Time flies in the presence of angry faces"... depending on the temporal task used! *Acta Psychologica, 136*, 354-362

Gresham, F. M. & Elliott, S. (2008). *Social Skills Improvement System*. Toronto: Pearson.

Grondin, S. (2020). *The Perception of Time – Your Questions Answered.* New York: Routledge.

Grondin, S., Laflamme, V., Bienvenue, P., Labonté, K. & Roy, M.-L. (2015). Sex effect in the temporal perception of faces expressing anger and shame. *International Journal of Comparative Psychology*, *28*.

Mioni, G., Stablum, F., McClintock, S. M., & Grondin, S. (2014). Different methods for reproducing time, different results. *Attention, Perception & Psychophysics*. *76,* 675-681.

Ouellet, C., Tétreault, É., & Grondin, S. (in press). Politically biased time perception and perspective. *Timing & Time Perception.*

Rutter, M., Bailey, A., & Lord, C. (2003). *The Social Communication Questionnaire*. Los Angeles: Western Psychological Services.

Tehrani-Doost, M., Noorazar, G., Shahrivar, Z., Khorrami Banaraki, A., Farhad Beigi, P., & Noorian, N. (2017). Is emotion recognition related to core symptoms of childhood ADHD? *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 26,* 31-38.

Toplak, M. E., Dockstader, C., & Tannock, R. (2006). Temporal information processing in ADHD: Findings to date and new methods. *Journal of Neuroscience Methods, 151*, 15-29.

Walg, M., Oepen, J., & Prior, H. (2015). Adjustment of time perception in the range of seconds and milliseconds: The nature of time processing alterations in children with ADHD. *Journal of Attention Disorders, 19*, 755-763.

Yuill, N. & Lyon, J. (2007). Selective difficulty in recognising facial expressions of emotion in boys with ADHD. *European Child and Adolescent Psychiatry, 16*, 398-404.

THE ENDLESS GLOBAL PRECEDENCE EFFECT: THE BEHAVIOURAL DYNAMICS BETWEEN GLOBAL PRECEDENCE AND FAST-SAME EFFECT

Aleksandar Bulajić[1, 2], Frouke Hermens[3], Sven Panis[1], Cees van Leeuwen[1, 4], and Thomas Lachmann[1, 4, 5]

[1]*Center for Cognitive Science, Cognitive and Developmental Psychology unit, University of Kaiserslautern, Kaiserslautern, Germany*

[2]*Andragogy Unit, University of Belgrade, Belgrade, Serbia*

[3]*Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, The Netherlands*

[4]*Laboratory for Perceptual Dynamics, Brain & Cognition Research Unit, University of Leuven, Leuven, Belgium*

[5]*Facultad de Lenguas y Educación de la Universidad Antonio de Nebrija, Madrid, Spain*

Abstract

The Global Precedence Effect (GPE) observed in compound figures (Navon stimuli) represents a global pattern (global shape) made up of local components (local elements). The perceptual identity of local elements with the global shape they compose (based on main gestalt principles) can result in either *congruent* or *incongruent* condition. The GPE is comprised of two effects: the Global Advantage Effect or GAE (overall faster and/or more accurate responses observed at the global level of attenuation) and Global Interference Effect GIE (overall faster and/or more accurate responses observed for the congruent condition, exclusively at the local level of attenuation). Although it is in general a robust and reliable effect, the GPE depends on a number of factors such as number of components, eccentricity, relative size, spatial frequency contrast, the meaningfulness of the stimuli and individual differences etc. Bearing those in mind, one may pose a question whether the effect involves processing of categorical representations. The aim of the current study was to investigate the *involvement of categorical representations in GPE* by using the *same-different* task. The typically utilized task that shows GPE is the *target identification task*, observed at local or global attenuation level of a single compound figure. Recently, a number of studies used compound stimuli in successive or simultaneous *same-different* tasks, without systematic investigation of the effects emerging from the interaction of the two paradigms. This type of task, allows for conditions in which *same-different* comparison can be perceived (physical level) or inferred (categorical level) on relevant and irrelevant attenuation level simultaneously – given that attention resources are allocated disproportionally over task relevant and irrelevant level. Moreover, potentially finding a *fast-same effect* interaction with GPE subeffects may give more clues on the role categorical representations have in GPE.

A total of 24 participants aged between 20 and 31 years took part in the study (12 male). We employed 2x2x2 repeated measures design encompassing the following factors: Instruction level (‘global’, ‘local’), Response type on the relevant task level (‘same’, ‘different’), Matching of the response type between relevant and irrelevant level (‘match’, ‘non-match’). By utilizing ANOVA and Event History Analysis (EHA) several important observations were found. Given that the overall accuracy was 97.9% and no trade-off with speed took place, we found following. A strong GAE main effect was observed – global level responses were always faster than local ones (additionally confirmed by post-hoc analyses), as well as a *fast-same* main effect (‘same’ faster than ‘different’ responses) and an overall significant main effect of Matching (faster responses when Response type [‘same’, ‘different’] is matched between relevant and irrelevant task level). The post-hoc analyses of interactions, especially the triple interaction found between all of the main effects revealed that the *same-fast* effect is not universal – while stable at the

global level it is partially cancelled at the local level. This particular interaction was moderated by the Matching factor, which gives a categorical dimension to stimuli-to-be-compared. For example, global-same-match condition represents a physical identity comparison, while global-same-nonmatch is a categorical identity one. The latter is implying the answer to our research question, i.e., that there is *categorical processing* involved in GPE given that: (i) it is only attributable to GIE, (ii) GIE depends on the complexity of visual scenery and/or special (task) requirement imposed by the percept.

EFECT OF STIMULUS PRESENTATION ON PITCH RANKING IN COCHLEAR IMPLANT USERS AND NORMAL-HEARING LISTENERS

Carmen M. Castañeda González[1], Werner Hemmert[1], and Sonja Karg[2]
[1]Munich Institute of Biomedical Engineering, Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany
[2]Medizintechnik und Biosignalverarbeitung, Fakultät Elektro- und Informationstechnik, Technische Hochschule Ingolstadt, Ingolstadt, Germany
<carmen.castaneda@tum.de>

Abstract

In cochlear implants (CIs), pitch can be conveyed via the selection of the electrode and thus via the tonotopy principle as in normal hearing. However, due to the current spread in the cochlea resulting in channel crosstalk, accurate pitch representation remains limited, making pitch perception one of the most difficult tasks for CI users. During a pitch ranking procedure, the subject decides which of two successive stimuli was higher in pitch. In our CI-study (Karg et al., 2018) (13 ears, 10 MED-EL users; ages: 22 to 55 years, mean: 44 years; direct stimulation via the RIB2), each of the 12 CI-electrodes was compared with its three next basal neighbors. Per electrode pair, each presentation order (descending, i.e., from basal to apical, and ascending, i.e. from apical to basal) was measured five times in random order. We observed that the presentation order of the electrodes had an influence on the pitch ranking's uncertainty. To investigate this presentation order effect, we did a comparable experiment with 8 normal-hearing (NH) subjects (ages: 22 to 28 years, mean: 25.5 years). The high channel crosstalk in electrical CI-stimulation was approximated by bandpass-filtered white Gaussian noise with shallow filter slopes (6 dB/oct). To investigate the effect of the slopes' steepness, we also tested a filter bank with steep slopes (60 dB/oct). NH subjects performed better with steep than with shallow filters: adjacent steep filters achieved 90.3 ± 8.4 % performance, adjacent shallow filters only 73.5 ± 13.4 %. We conclude that the difficulty level of the pitch ranking for CI users could be successfully approximated in NH subjects using the shallow filters. Therefore, pitch ranking performances were comparable between both groups (adjacent CI-electrodes: 73.7 ± 10.2 %) and improved with increasing distance between the electrodes (85.8 ± 9.1 %, 90.6 ± 9.0 %) and (shallow) filters (87.6 ± 13.5 %, 93.6 ± 7.9 %). In the pitch rankings of adjacent CI-electrodes and adjacent shallow filters, a Greenhouse-Geisser-corrected repeated measures ANOVA revealed a significant interaction between presentation order and stimulation site for CI users [$F(5.09,55.95) = 4.16$, $p = .003$, $\eta^2 = .27$] and NH subjects [$F(3.63,25.46) = 8.79$, $p < .001$, $\eta^2 = .56$]. Bonferroni-corrected post-hoc tests showed that in apical locations, descending pitch pairs reached significantly higher performances than ascending pairs in CI users and NH subjects. Conversely, ascending pitch pairs were recognized significantly better in basal regions in both groups. Our results suggest that CI users and NH subjects exhibit similar directional effects in difficult pitch perception tasks, the cause of which is subject to further investigation.

References

Karg, S., Huber, M., Hemmert, W., & Völk, F. (2018). Kriterien von Cochlea-Implantat-Nutzern zur Beurteilung der Ratentonhöhe. *Fortschritte der Akustik*, DAGA 2018.

VISUAL ATTENTION GUIDED BY EITHER ASSOCIATIVELY OR PROPOSTIONALLY LEARNED VALENCE

Alexandra E Clausen[1], Hannah Lindenkreuz[1], Henri Kracke[1], and Florian Kattner[2]
[1]Department of Psychology, Technical University Darmstadt
[2]Department of Psychology, HMU Potsdam

Abstract

The learning processes involved in evaluative conditioning (EC) has been a topic of discussion for more than four decades. EC refers to the change in valence of a formerly neutral stimulus after having been paired with an affective unconditioned stimulus (US). The present experiment is an attempt to contribute to the discussion of propositional and associative learning mechanisms involved in EC. Therefore, we tested whether (1) different processes during learning (propositional vs. associative) produce similar magnitudes of EC, and (2) visual attention is influenced more by associative than by propositional EC. In a mixed design participants were randomly assigned to either a group that had to respond quickly to the US (fostering automatic associative learning) or a group that was asked to predict the valence (positive, negative, or neutral) of each US (propositional learning). We expect that evaluative ratings change as a function of the valence of the US after both associative and propositional EC, but contingency awareness is expected to be higher after propositional learning. In addition, an associatively acquired valence is expected to be reflected more strongly in visual attention, as measured with a visual search task with the conditioned stimulus as the target.

THE POSITIVE EFFECT OF ALCOHOL CONSUMPTION ON VISUAL ABSOLUTE THRESHOLD

Evyatar Cohen and Yaniv Mama.
Ariel University, Ariel, Israel

Abstract

The cognitive filter (Broadbent, 1958) is the gate to awareness. Stimulus which does not reach minimal intensity will not pass the filter and will not be detected. Inhibiting this filter should result in lowering absolute threshold. Indeed, former studies found that Low-Dose Alcohol (LDA) consumption influenced the absolute threshold for olfactory stimuli (Endevelt-Shapira et al., 2014). In the current work, LDA consumption was used to inhibit the function of the cognitive filter with visual stimuli. The results depict lower threshold for visual stimuli after LDA consumption relative to no-alcohol condition. These results, together with previous studies using olfactory stimuli and auditory stimuli (Murata, 2011), suggest a common denominator – the cognitive filter – as the basis for this effect. Here we offer a new model that puts the cognitive filter in the heart of threshold detection. Moreover, we introduce a new mean to influence it.

Often, the influence of alcohol on our senses is perceived negatively, such that alcohol is thought to lead to impaired sensory perception. However, studies (Endevelt-Shapira, et al., 2014) that have used the absolute threshold paradigm to examine the influence of low-dose alcohol (LDA) consumption have shown that it may lead to improved detection of various stimuli via the sense of smell. The authors suggested that a physiological account, i.e., LDA affect the brain in a way that it enhances one's sense of smell by promoting disinhibition of the olfactory bulb in the frontal lobe (Endevelt-Shapira, et al., 2014). Hence, small portion of alcohol may result in a lower sensory absolute threshold (i.e., a weaker stimulus is needed in order to elicit a response).

In the current study, in line with Broadbent's filter model (Broadbent, 1958), we argue that alcohol reduces the effect of the cognitive filter, which is responsible for filtering stimuli from the stage in which they are perceived at a sensory level to the stage in which they are perceived at the cognitive level. The cognitive filter theory describes the way in which we perceive stimuli and explains that we do not consciously and cognitively perceive all the stimuli that we observe at the neural level. This discrepancy enables us to manage stimuli more effectively and allows us to pay attention to the more relevant stimuli in our environment (Lachter, Forster, & Ruthruff, 2004; Wickens, 2021).

To sum up, we assume that alcohol inhibits the cognitive filter, which leads to an increase in the ability to detect stimulus which is not strong enough to cross the filter in a typical (non-alcohol) state. Here, this hypothesis was tested with regards to the sense of sight by using visual stimuli. To the extent that alcohol has a similar effect on our sense of sight as it does on smell, we can determine that it is not due only to a physiological effect on the sense of smell, but rather a more general, cognitive phenomenon based on the cognitive filter.

Method

In the current experiment, we used a modified version of the staircase method for the absolute threshold (Endevelt-Shapira et al., 2014; Experiment 1). In Experiment 1, the researchers tried to show that a low dose of alcohol may have a positive effect on participants who were asked to identify the smell of a rose (the smell of Phenyl Ethyl Alcohol – PEA – which has a rosy smell). To test that, they used the staircase paradigm to detect the absolute threshold. Each participant participated in the alcohol condition and the no-alcohol condition (a within subject design). The alcohol condition contained a grape juice (120 ml) mixed with vodka (35 ml) beverage, as opposed to the condition without alcohol, with only grape juice beverage (155 ml). After drinking the beverage, the participants were asked to detect a smell of a rose from 3 different jars, 2 of them did not carry any smell, and 1 contained the smell of a rose (PEA). As the participant was unable to identify the stimuli, the concentration was increased. After the participant was able to identify the smell correctly, and had two consecutive HITS, the concentration was decreased to the point that the participant no longer was able to identify it. The results pointed to a positive significant effect for the alcohol condition.

In the current study, we used the same conditions that were used in Endevelt-Shapira et al (2014) but instead of olfactory stimuli we used visual. Upon arrival to the lab, participants were randomly assigned to one of the two conditions (alcohol/noalcohol) and were asked to consume the specified drink for that condition and to complete a demographic questionnaire (participants came in two different occasions in order to participate in both conditions, order was counterbalanced). Five minutes after consuming the drink, participants began the visual experiment phase. During the experiment, participants were asked to detect the presence of a figure of a rose (Figure 1) on a computer screen between 3 locations (left, middle, right). The distance from the screen was set to 45 cm.

Fig 1. Example of the image that was present on the screen. The level of brightness was changed according to the staircase method, and the location was changed randomly.

In each step, the screen was divided into three parts, and a rose of varying levels of brightness always appeared in one of the locations. The location of the rose varied randomly after each keypress. Participants were asked to press a key on the keyboard in correspondence with the location of the stimuli that they had detected. If no stimulus was detected, participants were asked to press any key on the keyboard. The brightness level of the rose presented varied according to the detection or non-detection reported by the participants and in accordance with the aforementioned staircase method, as the participants were exposed to 14 ascending and 14 descending sequences of stimulus intensity levels. Each instance of stimulus detection was verified twice to avoid a situation in which a participant would move up a step due to an accidental press of the correct button. After data collection was complete, we calculated the average of the last four sequences for each participant across both experimental conditions and compared the results of each condition by the participant.

Results and Discussion

To test whether there was a difference in the absolute threshold in the visual detection task after LDA consumption versus the non-alcohol condition, we performed a dependent t-test. We found that there was a significant difference between the conditions, to the decree that when participants consumed a LDA, their detection threshold was lower (M = 39.15; SD = 0.51; luminosity 0.98) than consumption of a non-alcohol beverage (M = 38.47; SD = 1.84; luminosity 1.19), $t(34) = 2.17$, $p < .05$. In other words, participants exhibited a higher absolute threshold in the non-alcohol condition compared to the LDA condition.

The results of the current experiment and findings from previous experiments conducted in this field (Endevelt-Shapira, et al., 2014; Murata, et al., 2001, who found similar results with auditory stimuli), in which participants are exposed to various stimuli after LDA consumption vs. no alcohol and then the absolute threshold is examined, we suggest that LDA consumption influences the absolute threshold in such a manner that it makes it more sensitive. In other words, LDA consumption increases the range of stimuli that participants can detect. Due to the fact that the current study replicated previous results using visual stimuli, we suggest that this phenomenon is not entirely explained by the physiology of one sense, but rather that there is a distinct cognitive explanation for this phenomenon as it occurs in a similar fashion across multiple senses as was found also in auditory stimuli (Murata, et al., 2001) and in the current work.

It is important to note that this “positive” effect of improving absolute threshold is the result of a “negative” effect of LDA on the filtering system. This cognitive filter enables the proper functioning of the attentional system. Lowering the absolute threshold through inhibition of this filter may result in allegedly better sensory abilities but it may also negatively affect related cognitive performance (e.g., Endevelt-Shapira and her colleagues also showed poorer performance in the Stroop paradigm after alcohol consumption; Endevelt-Shapira, et al., 2014).

The results of the current work sheds new light on one of the most well-known models in cognitive psychology. The effect of LDA consumption on the cognitive filter opens a new opportunity for exploring and manipulating this filter. To recap, LDA consumption lowers the absolute threshold such that we are able to detect weaker stimuli (i.e., it enters our consciousness) than when we are in a typical (no alcohol) state. Simply put- drink a little, you will be more sensitive.

Acknowledgements

We wish to thank Yuval Himai and Gaya Almkais for their assistance.

References

Broadbent, D. E. (1958). *Perception and communication*. Pergamon Press.

Cohn, T. E. (1981). Absolute threshold: Analysis in terms of uncertainty. *Journal of the Optical Society of America*, *71*(6), 783-785.

Endevelt-Shapira, Y., Shushan, S., Roth, Y., & Sobel, N. (2014). Disinhibition of olfaction: Human olfactory performance improves following low levels of alcohol. *Behavioural Brain Research, 272*, 66-74.

Murata, K., Kawashima, M., & Inaba, R. (2001). Auditory threshold reduction on alcohol ingestion. *Psychopharmacology, 157*(2), 188-192.

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review*, *111*(4), 880-913.

Potenza, M. N. & De Wit, H. (2010). Control yourself: alcohol and impulsivity. *Alcoholism: Clinical and Experimental Research, 34*(8), 1303- 1305.

Wickens, C. (2021). Attention: Theory, principles, models and applications. *International Journal of Human–Computer Interaction, 37*(5), 403417.

CHECKERBOARD SPEECH: A TROUGH IN THE INTELLIGIBILITY CURVES

Linh Doan[1], Kazuo Ueda[2], Hiroshige Takeichi[3], and Gerard B. Remijn[4]
[1]Human Science Course, Graduate School of Design, Kyushu University, Fukuoka 815-8540, Japan
[2]Department of Acoustic Design/Research Center for Applied Perceptual Science/Research and Development Center for Five-Sense Devices, Kyushu University, Fukuoka 815-8540, Japan
[3]Open Systems Information Science Team, Advanced Data Science Project (ADSP), RIKEN, Information R&D and Strategy Headquarters (R-IH), RIKEN, Kanagawa 230-0045, Japan
[4]Department of Acoustic Design/Research Center for Applied Perceptual Science, Kyushu University, Fukuoka 815-8540, Japan
<doan.linh.298@s.kyushu-u.ac.jp; ueda@design.kyushu-u.ac.jp; takeichi@riken.jp; remijn@design.kyushu-u.ac.jp>

Abstract

Ueda, Kawakami, and Takeichi [(2021). JASA Express Lett. 1, 075204] investigated the intelligibility of checkerboard speech stimuli (segmented and interrupted in frequency and time) in comparison with the intelligibility of interrupted speech stimuli. They showed that the intelligibility of the checkerboard speech stimuli with 20 frequency bands (critical bands) was almost 100% irrespective of the segment duration (20-320 ms), but that the intelligibility of the checkerboard speech stimuli with 2 and 4 frequency bands (based on the factor analysis results for 8 languages/dialects by Ueda and Nakajima [(2017). Sci. Rep. 7, 42468]) was always lower than the intelligibility of the interrupted speech stimuli, with the minimum intelligibility of about 40% at the 160-ms segment duration. However, in the previous study, only three steps of frequency bands (2, 4, and 20) were utilized. Therefore, the goal of this study is (1) to replicate the previous experiment while extending their approach, (2) to have a closer look at how the number of frequency bands and segment durations affect the intelligibility of checkerboard speech stimuli. Current experiment 1 included 8- and 16-band stimuli, in which the 4-band stimuli in the previous study were subdivided, yielding four steps of frequency bands (4, 8, 16, and 20). Current experiment 2 included 4- and 8-band stimuli, which were divided in equal steps in critical bands, also yielding four steps of frequency bands (4, 8, 4-critical, 8-critical, 20). In both experiments, the intelligibility of the interrupted speech stimuli declined from approximately 95% at the 20-ms segment duration to 55% at the 320-ms segment duration. In experiment 1 ($n = 20$), the results demonstrated that the intelligibility for the 20- and 16-band checkerboard speech stimuli were close to perfect (100%) regardless of the segment duration. The intelligibility of the 8-band stimuli was lowest (84%) at the 160-ms segment duration, and was always better than the intelligibility of the interrupted speech stimuli with the same segment duration. In experiment 2 ($n = 20$), the ways of dividing frequency bands basically unaffected the intelligibility of the 4- and 8-band stimuli. The results suggest that the trough of the intelligibility curves may be related to the limits of auditory grouping for the checkerboard speech stimuli, and that the intelligibility of the checkerboard speech stimuli may be subject to how far the 4 frequency bands obtained from the factor analysis results are included at each moment.

AGE-RELATED DIFFERENCES IN THE PERCEPTION OF SPOKEN EMOTIONS: THE PSYCHOPHYSICS OF EMOTIONAL SPEECH-IN-NOISE PROCCESING

Yehuda Dor[1, 2], Vered Shakuf[3], Boaz M. Ben-David[2, 4, 5], and Daniel Algom[1]
[1]*School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel*
[2]*Communication, Aging and Neuropsychology Lab (CANlab), Baruch Ivcher School of Psychology, Reichman University (IDC), Herzliya, Israel*
[3]*Department of Communications Disorders, Achva Academic College, Arugot, Israel*
[4]*Toronto Rehabilitation Institute, University Health Networks (UHN), Toronto, ON, Canada*
[5]*Department of Speech-Language Pathology, University of Toronto, Toronto, ON, Canada*

Abstract

Older adults perceive emotions in speech differently than young adults. Evidence in the literature suggest that these differences can be attributed in part to age-related sensory changes. However, it remains unclear how these changes affect the processing of emotional information in the semantic (meaning of the words) and prosodic (tone) speech channels, as well as the integration of these speech channels. To answer these questions, in the current study young and older listeners preformed the Test for Rating of Emotions in speech (T-RES), comprising spoken sentences carrying different combinations of emotions in prosody and semantics. Stimuli were presented on the background of masking speech-spectrum noise or 8-talker babble noise (SNR=-4 dB for young adults; 0 dB for older adults), and performance was compared to that in quiet. Age-related differences in performance in different listening conditions highlighted the role of sensory factors in the processing of emotional speech.

INDUCING EXPECTATIONS WITH STAR RATINGS INFLUENCES THE APPRECIATION OF SHORT MOVIES

Joanie Drouin, Pier-Alexandre Rioux, Daniel Fortin-Guichard, and Simon Grondin
Université Laval, Québec, Canada

Abstract

In this experiment, the expectation effect was tested on two movies. In a no expectation condition, two groups of participants assigned a high mean score to one movie (positive movie) and a low mean score to another (negative movie). In four other groups of 20 participants, they were given a fictitious rating of 8.5 or 4.5 (out of 10) before viewing either the positive or the negative movie. Results show that both movies received higher scores when associated with fictitious positive scores.

Positive or negative expectations are formed when a person has access to information describing a stimulus (Zellner et al., 2004). For example, when a product is rated 4 out of 5 stars by other consumers, expectations tend to be positive, which can influence appreciation (Furnham & Chu Boo, 2011). An "assimilation effect" is observed when positive expectations lead to higher rating scores or when negative expectations tend to decrease appreciation (Davidenko et al., 2015; Zellner et al., 2004). Depending on the type of information, the processing differs. It is possible to use a rapid, but biased process called heuristic processing, as well as a detailed, systematic process that requires a lot of cognitive effort (Chaiken, 1987; Yeomans et al., 2008).

Heuristic processing, among others, occurs when participants must evaluate a product that is associated with a rating score (for example, 4 out of 5 stars). Since this rating already offers a point of comparison, no further analyses are needed to evaluate the product (Zellner et al., 2004). Participants engage in a heuristic treatment, which can be referred to as a "mental shortcut" (Chaiken, 1987). This type of treatment is fast because it requires little attention, effort, or cognitive resources (Chaiken, 1987; Guazzini et al., 2015). It is automatically activated and facilitates the decision-making process. However, such a shortcut can generate biases, like a cognitive error called the anchoring heuristic (Shiloh et al., 2002). This error occurs precisely when participants have access to ratings given by other individuals. They tend to rely on the available information, which creates positive or negative expectations, and they try to find similarities with their own opinion. Participants quickly adapt their responses according to the valence of the available rating (Cardello & Sawyer, 1992; Chapman & Johnson, 1999; Wilson et al., 1989; Zuckerman & Chaiken, 1998). This procedure is believed to explain the assimilation effect (Zellner et al., 2004), which seems to be observed when no correct answer is expected (Onuma et al., 2013; Robinson et al., 2007).

On the other hand, a systematic treatment occurs when participants must evaluate a stimulus presented without any clear information: there are no ratings or adjectives that could give an idea about the product value. Sometimes, only the name of the product is available. That is sufficient to generate positive or negative expectations, but since they are not a reliable source of information, participants have no point of comparison (Zellner et al., 2004). A detailed systematic treatment is required to evaluate the product, which takes more time and cognitive resources (Chaiken, 1987). Any difference between expectations and the real value of the stimulus is analyzed. Most of the time, because expectations are uncertain, participants with positive expectations find the product disappointing. A negative hedonic contrast occurs between their expectations and their appreciation scores. A positive hedonic contrast can be observed when the stimulus is more appreciated than expected because of low expectations (Zellner et al., 2004).

Usually, hedonic contrasts appear when a positive stimulus receives greater appreciation scores when presented after a negative one, and vice-versa. Studies have shown the presence of hedonic contrasts with artistic stimuli such as paintings (Zellner et al. 2010), music (Parker et al., 2008), and most recently short movies (Drouin et al., 2022). As reported earlier, expectations can also create hedonic contrasts. Mechanisms explaining why expectations result in assimilation effects or hedonic contrasts remain unclear, especially with complex stimuli such as movies.

In the present study, a fictitious rating value, positive or negative, was assigned to each movie and reported to participants before the viewing. We hypothesized that assimilation effects would occur when both movies are associated with negative rating scores, so participants would give lower ratings in these conditions (Zellner et al., 2004). Based on the work of Zellner et al. (2004), a second hypothesis could be that assimilation effects also occur when both movies are associated with positive ratings. However, we think that hedonic contrasts, and not an assimilation effect, will be observed, even though ratings are normally associated with assimilation. We believe that positive expectations could create a strong disappointment among the participants, so they would give lower ratings.

Method

Pilot project

Two animated short movies, lasting about 4 minutes, were initially chosen by the first author: "Crooked Rot" (https://www.youtube.com/watch?v=oYjny4qNy24,_Firth, 2008) and "Dji. Death Fails" (https://www.youtube.com/watch?v=kkxNaJ2q4QA, Voloshin & Djumaev, 2012). In a pilot project, 24 people, 10 women and 14 men (*M* age = 26.41 years), viewed one of the two movies. Following the viewing, each participant had to write a rating score of 1 to 10 (10 = the movie was much appreciated) on a questionnaire. Each participant involved in the pilot project was installed alone in a quiet room of the Psychology of Perception Research Laboratory, at Université Laval.

The results of the pilot project revealed that "Crooked Rot" received an average score of 4.5, with scores ranging from 2 to 7. "Dji. Death Fails" received an average score of 7.4, the ratings varying from 5 to 9. The difference between the rating scores is significant, $t(22) = 4.24$, $p < .05$. Thus, for the study, "Crooked Rot" and "Dji. Death Fails" were adopted as the negative and positive movies, respectively.

Participants

In this study, 120 participants, 74 women and 46 men (M age = 27.13), were recruited from Université Laval and randomly assigned to one of six groups. To avoid appreciation biases, they had no uncorrected vision or hearing problems. They had no diagnosed neurological or psychopathological disorders. All participants signed a consent form and agreed freely to participate in the experiment. As compensation for their time, they had the chance to win one of four 25$ gift certificates for a movie theatre in Quebec City.

Groups 1 and 2. In order to verify the presence of hedonic contrasts, participants of Group 1 first viewed the positive movie "Dji. Death Fails", then the negative movie "Crooked Rot", and participants of Group 2 watched both movies in the reverse order. No description or instruction was given to the participants of these groups.

Group 3. Participants of this group only viewed the positive movie. They read this description: “The short movie that you will see is called ‘Dji. Death Fails’ and was directed by Dmitry Voloshin. It received positive ratings of 8.2 / 10 from Rotten Tomatoes and 8.5 / 10 from IMDb. The ratings of these two cinematic review sites are based on the average appreciation score given by many movie lovers”.

Group 4. Participants of this group only viewed the positive movie. The same text as the one used with Group 3 was read, but this time, the movie was associated with negative ratings of 4.2 and 4.5 out of 10.

Group 5. Participants of this group only watched the negative movie. Before the viewing, they read this description: “The short movie you will see is called ‘Crooked Rot’ and was directed by David Firth. It received positive ratings of 8.2 / 10 from Rotten Tomatoes and 8.5 / 10 from IMDb. The ratings of these two cinematic review sites are based on the average appreciation score given by many movie lovers”.

Group 6. Participants of this group only watched the negative movie. The same text as the one used with Group 5 was read, but this time, the movie was associated with negative ratings of 4.2 and 4.5 out of 10.

Procedure

All participants viewed the short movies on a Lenovo ThinkVision brand computer monitor, model 7360PC7. The computer screen was 34.3 cm high by 36 cm wide and had a resolution of 1055 x 980.

After each movie presentation, all participants completed a home assessment questionnaire to indicate, on a scale of 1 to 10, their level of appreciation (1 = the movie was considered very bad, 10 = the movie was considered very good). Moreover, after the experiment, they completed a homemade demographic questionnaire that gave information about their gender, age, field of study, and their level of knowledge about cinema. All these factors were controlled because they could influence movie appreciation.

Results

Table 2 shows the mean and standard error of the scores obtained in Groups 1, 2, 3, 4, 5, and 6. The rating scores of the first movie presented in Groups 1 and 2 (of Experiment 1) were included in the analysis as a control (no expectation) group. A 3 (Expectations: None vs. Negative vs. Positive) × 2 (Movie: Positive vs. Negative) factorial designed ANOVA was performed on the rating scores. Results reveal that the main effect of the Movie is statistically significant, $F(1, 113) = 46.07$, $p < 0.001$, $\eta^2_p = 0.290$. Participants therefore assigned higher scores to the positive (M = 6.81) than to the negative movie (M = 4.52). In addition, the main effect of the Expectations was significant, $F(2, 113) = 4.12$, $p = 0.019$, $\eta^2_p = 0.068$. Multiple comparison tests showed that groups (3 and 5) with positive expectations had significantly higher scores (Positive movie: M = 7.45; Negative movie: M = 5.25) than groups (1 and 2) with no expectations (Positive movie: M = 7; Negative movie: M = 3.60), $t(78) = 2.13$, $p = 0.029$, d = 0.480. However, the score of the groups (4 and 6) with negative expectations (Positive movie: 6.05; Negative movie: M = 4.07) was neither significantly different from the score of the groups with positive expectations, $t(79) = -2.09$, $p = .040$, $p = .052$, d = -0.467, nor from the scores of the groups without expectations, $t(79) = 0.29$, $p = .968$, d = 0.407. Finally, the Expectation × Movie interaction was not statistically significant, $F(2, 113) = 2.90$, $p = .059$, $\eta^2_p = 0.049$.

Table 1. Mean Appreciation Scores (and Standard Error), in Each Experimental Group for both Movies

Group	Movie	
	Positive	Negative
1. Positive movie – No expectation	7 (1.86)	-
2. Negative movie – No expectation	-	3.60 (1.57)
3. Positive movie only – Positive expectation	7.45 (1.57)	-
4. Positive movie only – Negative expectation	6.05 (2.41)	
5. Negative movie only – Positive expectation	-	5.25 (2.27)
6. Negative movie only – Negative expectation	-	4.07 (2.29)

Discussion

The objective of this study was to test the effect of expectations on movie appreciation. Results revealed that changing expectations influences appreciation. Contrary to our hypothesis, both positive and negative movies received higher ratings when participants had access to positive reviews. Hedonic contrasts resulting in strong disappointment was not observed. It seems that assimilation effects (Zellner et al., 2004) occurred for both movies, but for different reasons.

For the positive movie, it is possible that participants' positive expectations were quickly confirmed: the first few seconds of this movie were sufficient to appreciate it, as the available positive ratings suggest. Ratings could have become reliable sources of information. In such a case, no further analyses were necessary. Participants may have used a mental shortcut (Chaiken, 1987), and relied on the false positive ratings, thus positively influencing their appreciation. This heuristic way to process information, and the concept of anchoring bias, may be responsible for the assimilation effect observed when positive ratings were associated with the positive movie. This is supported by the fact that participants had no correct response to give and, therefore, no social pressure, and that the real estimated value of this movie, as determined by the pilot study, was positive. However, it is important to note that in the present study, there was no method to distinguish the type of processing (systematic or heuristic) adopted by the participants. In future research, the method should control for cognitive load, available time, and motivation, since these factors could determine the type of treatment adopted by participants (Furnham & Chu Boo, 2011).

As for the negative movie with positive expectations, another phenomenon could explain the assimilation effect. Participants probably did not find that their positive expectations were congruent with the negative impression left by this dark and bizarre movie. It is possible that this short movie is too confusing for a solid opinion to emerge, contrary to the positive one. Several participants reported that they did not understand the meaning of the work. Thus, by seeing positive scores and realizing their strange impression about the movie, they could have doubted their judgment. Positive ratings may have contradicted their negative opinion, which may have led to a cognitive dissonance and a certain degree of discomfort (Festinger & Carlsmith, 1959; Hinojosa et al., 2016). Participants could have modified their beliefs and thought they did not understand this movie enough to be sure of their opinion. They probably gave higher ratings by relying on the positive scores available. This explanation is supported by the fact that most participants did not describe themselves as movie experts in the demographic questionnaire.

In addition, in an attempt to give meaning to the negative movie, participants could have used a systematic way of processing information, which requires more cognitive resources (Chaiken, 1987). By focusing on certain details (e.g., "Stop Motion" animation technique, sound effects) and by taking more time to analyze them, appreciation could be positively influenced. Moreover, Schultze et al. (2012) noted that to reduce the discomfort caused by cognitive dissonance, individuals tend to select the information corresponding to the most optimal response. According to Zellner and collaborators (2004), when ratings are associated with a stimulus, assimilation effect is likely to occur. However, generally speaking, systematic treatment is associated with unknown stimuli (without ratings) and with hedonic contrasts. Although ratings seemed to create assimilation effects, the negative (and bizarre) movie could remain sufficiently ambiguous in the eyes of the participants to require detailed processing of the information.

Acknowledgements

This study was supported by a scholarship from the Social Sciences and Humanities Research Council (SSHRC) of Canada to JD and by a research grant from the SSHRC to SG.

References

Cardello, A. V. & Sawyer, F. M. (1992). Effects of disconfirmed consumer expectations on food acceptability. *Journal of Sensory Studies, 4*, 253-277. doi: 10.1111/j.1745459X.1992.tb00194.x

Chaiken, S. (1987). The heuristic model of persuasion. In P. M. Zanna, J. M. Olson & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (vol. 5, pp. 3-39). Hillsdale, NJ: Erlbaum.

Chapman, G. B. & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes, 79*(2), 1-39. doi:10.1006/obhd.1999.2841

Davidenko, O., Delarue, J., Marsset-Baglieri, A., Formentin, G., Tomé, D., Nadkarni, N., & Darcel, N. (2015). Assimilation and contrast are on the same scale of food anticipated-experienced pleasure divergence. *Appetite, 90*(1), 160-167. doi : 10.1016/j.appet.2015.03.006

Drouin, J., Racine, C., Savard, C., & Grondin, S. (2022). Effect of Hedonic Contrasts on Movie Appreciation. *Psychology of Aesthetics, Creativity, and the Arts*. In press. https://doi.org/10.1037/aca000048

Festinger, L. & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58*(2), 203-210.

Firth, D. (Producer and scenarist). (2008). *Crooked rot* [Short-movies]. United Kingdom.

Furnham, A. & Chu Boo, H. (2011). A literature review of the anchoring effect. *Journal of Socio-Economics, 40*(1), 35-42.

Guazzini, A., Yoneki, E., & Gronchi, G. (2015). Cognitive dissonance and social influence effects on preference judgments: An eye tracking based system for their automatic assessment. *International Journal of Human-Computer Studies, 73*(1), 12-18.

Hinojosa, A. S., Gardner, W. L., Walker, J., & Gullifor, D. (2016). A review of cognitive dissonance theory in management research: Opportunities for further development. *Journal of Management, 43*, 1080-1109. doi: 10.1177/0149206316668236

Onuma, T., Fuchimoto, J., & Sakai, N. (2013). Brand categorization and hedonic transfer: negative evaluations of a beverage transfer to a novel beverage from the same brand. *Tohoku Psychologica Folia, 72*(1), 46−62.

Parker, S., Bascom, J., Rabinovitz, B., & Zellner, D. A. (2008). Positive and negative hedonic contrast with musical stimuli. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(3), 171–174. doi:10.1037/1931-3896.2.3.171

Robinson, T. N., Borzekowski, D. L. G., Matheson, D. M. M., & Kraemer, H. C. (2007). Effects of fast food branding on young children's taste preferences. *Archives of Pediatrics & Adolescent Medicine, 161*(1), 792-797.

Shiloh, S., Salto, E. & Sharabi, D. (2002). Individual differences in rational and intuitive thinking styles as predictors of heuristic responses and framing effects. *Personality and Individual Differences, 32*(1), 415-429. doi: 10.1016/S0191-8869(01)00034-4

Schultze, T., Pfeiffer, F., & Schulz-Hardt, S. (2012). Biased information processing in the escalation paradigm: Information search and information evaluation as potential mediators of escalating commitment. *Journal of Applied Psychology, 97*(1), 16-32. doi: 10.1037/a0024739

Voloshin, D. (Producer) & Djumaev, S. (Scenarsit). (2012). *Dji. Death Fails* [Short-movies]. Moldavie.

Wilson, T. D., Lisle, D. J, Kraft, D., & Wetzel, C. G. (1989). Preferences as expectation-driven influences: Effects of affective expectations on affective experience. *Journal of Personality and Social Psychology, 56*(4), 519-530.

Yeomans, M. R., Chambers, L., Blumenthal, H., & Blake, A. (2008). The role of expectancy in sensory and hedonic evaluation: The case of smoked salmon ice-cream. *Food Quality and Preference, 19*(6), 565-573. doi: 10.1016/j.foodqual.2008.02.009

Zellner, D. A., Strickhouser, D., & Tornow, C.E. (2004). Disconfirmed hedonic expectations produce perceptual contrast, not assimilation. *American Journal of Psychology, 117*(3), 363-387.

Zellner, D. A., Jones, K., Morino, J., Cogan, E., Jennings, E., & Parker, S. (2010). Increased hedonic differences despite increases in hedonic range. *Attention, Perception, & Psychophysics, 72*(5), 1261-1265. doi : 10.3758/APP.72.5.1261

Zuckerman, A. & Chaiken, S. (1998). A heuristic-systematic processing analysis of the effectiveness of product warning labels. *Psychology & Marketing, 15*(7), 621-642.

RELIABILITY OF THE IRRELEVANT SPEECH EFFECT RECONSIDERED

Wolfgang Ellermeier[1] and Florian Kattner[2]
[1]Department of Psychology, Technical University Darmstadt
[2]Department of Psychology, Health and Medical University Potsdam
<*ellermeier@psychologie.tu-darmstadt.de, florian.kattner@health-and-medical-university.de*>

Abstract

The susceptibility to distraction by irrelevant sound differs considerably between individuals. In an article published some 25 years ago (Ellermeier & Zimmer, 1997), we showed that it remains relatively stable over time (test-retest reliability: r_{tt} = 0.45), and, when correcting for the low internal consistency of the trial-by-trial measurements, yields a stability coefficient as high as 0.82. The present study is the first to attempt to replicate this finding, while adding new experimental conditions not investigated in the original study, such as a steady-state distractor and irrelevant sound containing unexpected acoustical deviations. Furthermore, data were gathered for different samples, both online and in the laboratory. Initial analysis of a data collection still in progress indicates that the moderate reliability of the irrelevant-speech effect (contrasting performance in silence with changing-state speech) may be replicated. A more detailed analysis of the effects of the different distractor conditions, however, focusing on the changing-state effect, the steady-state effect, or the deviation effect, reveals test-retest correlations that are much lower. This pattern of outcomes will be interpreted on the basis of more recent theorizing about the irrelevant sound effect, particularly the duplex-mechanism account (Hughes et al., 2005, 2007).

EXPERIENCING ART IS BECOMING ART: UNDERSTANDING VIEWER BEHAVIOUR IN THE CONTEXT OF IMAGE FRACTALITY

Mark A. Elliott and Emma Kelly
School of Psychology, National University of Ireland Galway
<mark.elliott@nuigalway.ie>

Abstract

Both art theorists and artists themselves suggest that the experience of enjoying art does not derive solely from the act of "viewing art"; it also occurs through a more general physical embodiment of the artwork. Scientific studies have typically focussed upon the role perceptual factors such as scene composition contribute to aesthetic judgment, although thus far, no study has combined this approach with an examination of viewers' behaviour to explain aesthetic judgment. Here we show that a universally applicable measure of artistic composition, and a measure associated by some with the behaviour of the artist – the fractal dimensionality of the artwork – explains some aesthetic judgments when considered in combination with viewer behaviour: viewers were allowed to choose the position in viewing space at which they most appreciated a work of art. When expressed as a function of the fractal expansion of the artwork into the third dimension (i.e., into the space occupied by the viewer), this position explained judgments of the physical beauty of the artwork. The other aesthetic dimensions reported concerned the emotive value and intellectual quality of the work, both of which appeared unrelated to viewer position. The relationship between viewers' position, physical beauty and fractal dimensionality was statistically more significant when viewer position was referenced to a preferred subregion of the artistic image. These results indicate that aesthetic judgment relates to a whole-body embodiment of the artwork in the sense that judgments of physical beauty relate to the location chosen by the viewer for the purpose of aesthetic judgment when that location relates to the fractal expansion of the artistic composition. More controversially, this finding means that a viewer will experience the physical aesthetic in a work of art because, in positioning themselves relative to its fractal expansion, they will have become a part of the artistic composition they are viewing.

DOES THE DEGREE OF ABSTRACTION IN A VIDEO STIMULUS ALTER THE EXPERIENCE OF VECTION?

Xuanru Guo[1], Yuki Morimoto[1], Takeharu Seno[1], and Stephen Palmisano[2]
[1]*Faculty of Design, Kyushu University, Minami-ku, Fukuoka, 815-8540, Japan*
[2]*School of Psychology, University of Wollongong, Wollongong, NSW 2522, Australia*
<guoxuanru1101@gmail.com, morimoto@design.kyushu-u.ac.jp >

Abstract

Illusions of self-motion (vection) can be induced by purely visual motion. Evidence suggests that more naturalistic inducing stimuli tend to increase intensity of such experiences. However, this topic has received little examination. The current study made use of videos of camera motion recorded at two different locations (one being more, and the other being less, familiar). Computer graphics (CG) technology was then used to generate stimuli with four different degrees of abstraction (inversely proportional to the degree of naturalness). Results showed that vection intensity decreased with an increase in the abstraction level. Importantly, we also found an interaction between stimulus abstraction and location. These findings confirm that higher level (i.e., top-down factors) – such as the subjective naturalness and familiarity of the inducing stimulus – can have significant effects on the observer's experience of vection. This study further demonstrates that post-production manipulation of video content (via CG technology) can significantly alter perceptions of both abstraction and self-motion.

In everyday life, when you are on a stationary train and the train on the opposite line starts to move, you may mistakenly feel you are moving. This well-known example of vection is referred to as the "train illusion". Such illusions of self-motion can be induced in physically stationary observers by moving, or simulating the motion of, their surrounding environment (e.g., Brandt et al., 1973; Palmisano et al., 2015; Seno et al., 2017). This vection is now often used for various purposes in VR environments, amusement parks, and also in animation and movies (Tokunaga et al., 2016).

In vection research, the stimuli typically used to induce these visual illusions of self-motion illusion are drifting gratings and schematic patterns of global optical flow (Guo et al., 2021; Seno et al., 2017). However, other studies have used more naturalistic stimuli (e.g., motion displays based on real-world street scenes). They have shown that vection strength increases as the visual stimuli become more naturalistic (Riecke et al., 2006; Schulte-Pelkum et al., 2003). In order to compare the effects of different levels of naturalness on vection, these studies either scrambled the positions of elements in the original naturalistic images (using a mosaic method) (Riecke et al., 2006) or they altered material properties of the surfaces represented in these images (Morimoto et al., 2019). Computer graphics (CG) technology has rarely been used in the literature to generate visual stimuli with different levels of naturalness directly. Therefore, this study intends to use CG technology to generate visual stimuli with different levels of naturalness based on the distance of contour points in our video self-motion stimuli.

As CG technology has developed, vection scenes have become more popular (e.g., Seno et al., 2018; Tokunaga et al., 2016). CG technologies have been used in previous studies to create vection stimuli, modulate vection strength (e.g., by adding jitter or oscillations to CG optic flow - Bossard et al., 2020) and simulate camera or whole scene movement (Sato et al., 2020). Furthermore, CG technology can also be used to extract moving scene features and process them in a similar fashion to real stimuli (Su et al., 2016). Consider a video recording representing continuous camera motion through a 3-D scene. From the perspective of CG (Španel et al., 2006), the more abstract the graph and the smaller the triangle density, the less similar this video is to natural visual input (normally presented to a moving observer). Another factor that may affect the perceived naturalness of a visual motion stimulus is its familiarity (does it represent self-motion through a familiar or an unfamiliar environment). Therefore, this study investigated the vection induced by stimuli with different degrees of abstraction (generated by CG technology) and different degrees of familiarity - in order to better understand the influence of naturalness on vection intensity more directly and comprehensively.

Methods

Participants and Ethics statement

Seventeen college students (9 females and 8 males) volunteered to participate in this experiment. Their ages ranged from 23 to 31 (M = 25.53 SD = 2.45) years old. All had normal or corrected-to-normal vision and no reported history of either attention deficit disorder, mental illness, or brain damage. The experimental procedure was approved in advance by the Ethics Committee of Kyushu University. Participants were informed about the experimental stimuli and procedures prior to the experiment. We also obtained their informed consent prior to data collection.

Apparatus and Stimulus

Each of our experimental stimuli was presented on a Plasma display (3D Viera 65-inch, Panasonic, Japan, with 1920 × 1080 pixels resolution at a 60-Hz refresh rate) and controlled by a computer (ALIENWARE 17R1, Dell USA). The experimental software was developed, using MATLAB R2014a (Mathworks, Natick, MA), and PsychToolbox-3 (Brainard, 1997). Experiments were conducted in a darkened room. The viewing distance was approximately 57 cm. We used mobile phone (iPhone 13 Pro, Apple, 2021, USA) to record two 30 s videos during natural walking in two different areas (see Figure 1). In both cases, the spatial resolution was

Kyushu University Ohashi Campus

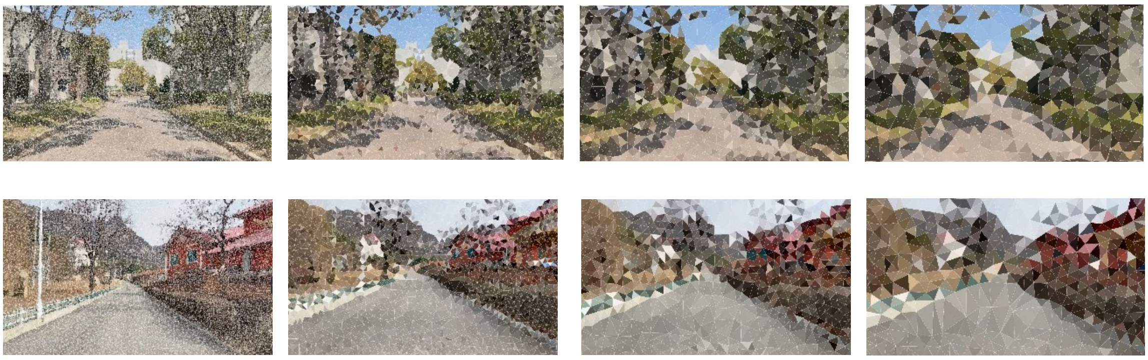
China Liaoning

Fig. 1. Original video sketch of two different areas.

1920 x 1080 pixels and the frame rate was 8 fps. We then used CG technology (Visual Studio 2017 OpenCV 450 C++; gray way; bilateral number: 2; sharpness: 0.2; low canny: 100, high canny: 200) to generate four new visual stimuli with different degrees of abstraction based on each of these two original videos (i.e., 8 new stimuli were created in total). The durations, frame rates, and spatial resolutions of these effect videos were the same as those in the original videos. The averaged luminance was always 30 cd/m2 and the luminance contrast was 0.6. Each of these video stimuli completely filled the screen of the Plasma display, which subtended a visual area of 100 degrees (horizontal) by 80 degrees (vertical).

Design and Procedure

The study had the following 4 x 2 within-subjects design: Video abstraction (which corresponded to the distance between contour points: Low 10; Medium to low 30; Medium to high 50; High 70 pixels) by video location (filmed either at the Kyushu University Ohashi Campus or in China Liaoning). Thus, there were 8 experimental stimulus conditions in total, as are shown in Figure 2.



Low Medium to low Medium to high High

Fig. 2. The four kinds of effect videos with different degrees of abstraction generated by CG technology. Top row: Kyushu University Ohashi Campus. Bottom row: China Liaoning.

The experimental procedure was as follows. Participants entered the laboratory and were asked to provide brief descriptions of both the naturalness (the degree to which it was close to the reality/nature) and the familiarity (whether they were familiar with the scene, it has nothing to do with practice) of each of the stimuli. They then watched the two original (unmodified) videos. As we recruited our participants from Kyushu University's Ohashi campus, they all indicated that they were more familiar with the videos from that campus (compared to those recorded in China). However, because there were individual differences in the frequency in which they attended the campus, we also had them rate each of these 8 video stimuli in terms of subjective familiarity. When a video stimulus was being presented, participants pressed the space bar on keyboard whenever they felt that they experienced self-motion, and kept it depressed as long as that vection experience continued. They were instructed that they should release this key during self-motion dropouts, and only press it again when their vection experience returned. The initial time of pressing the space bar was taken as the latency to vection onset, and the total time of pressing the space bar was taken as the vection duration. After each 30s exposure, a scoring interface appeared on the display so that participants could estimate the magnitude of their self-motion, the naturalness of the video stimulus, and the familiarity of the video stimulus shown in that trial, on scales ranging from 0 to 100. For vection magnitude ratings, 0 indicated “no vection was perceived” and 100 indicated “very strong vection was perceived”. For their naturalness ratings, 0 indicated “the video stimulus was very unnatural” and 100 indicated “video stimulus was very natural”. For their familiarity ratings, 0 indicated “the video stimulus was very unfamiliar” and 100 indicated “the video stimulus was very familiar”. Each condition was repeated four times throughout the experiment, which lasted approximately 20 minutes. There were four blocks of trials. At the end of each block participants could choose to either rest or continue.

Results and Discussion

Relationships between level of abstraction and vection intensity

We first calculated the latencies and durations of the vection under different each of the different stimulus conditions. We also sorted the subjective ratings of vection magnitude, naturalness and familiarity based on these conditions. These data were then averaged over four trial repetitions and analyzed statistically. As shown in Figure 3 below, the average vection magnitudes, latencies and durations all varied as a function of the level of video abstraction in each of video location conditions. Specifically, vection intensity can be seen to decrease as the level of video abstraction increased (as indicated by smaller magnitudes, longer latencies and shorter durations). In addition, vection intensity can be seen to increase for the more familiar Kyushu University campus videos (compared to the visual stimuli based on the video content filmed in China).

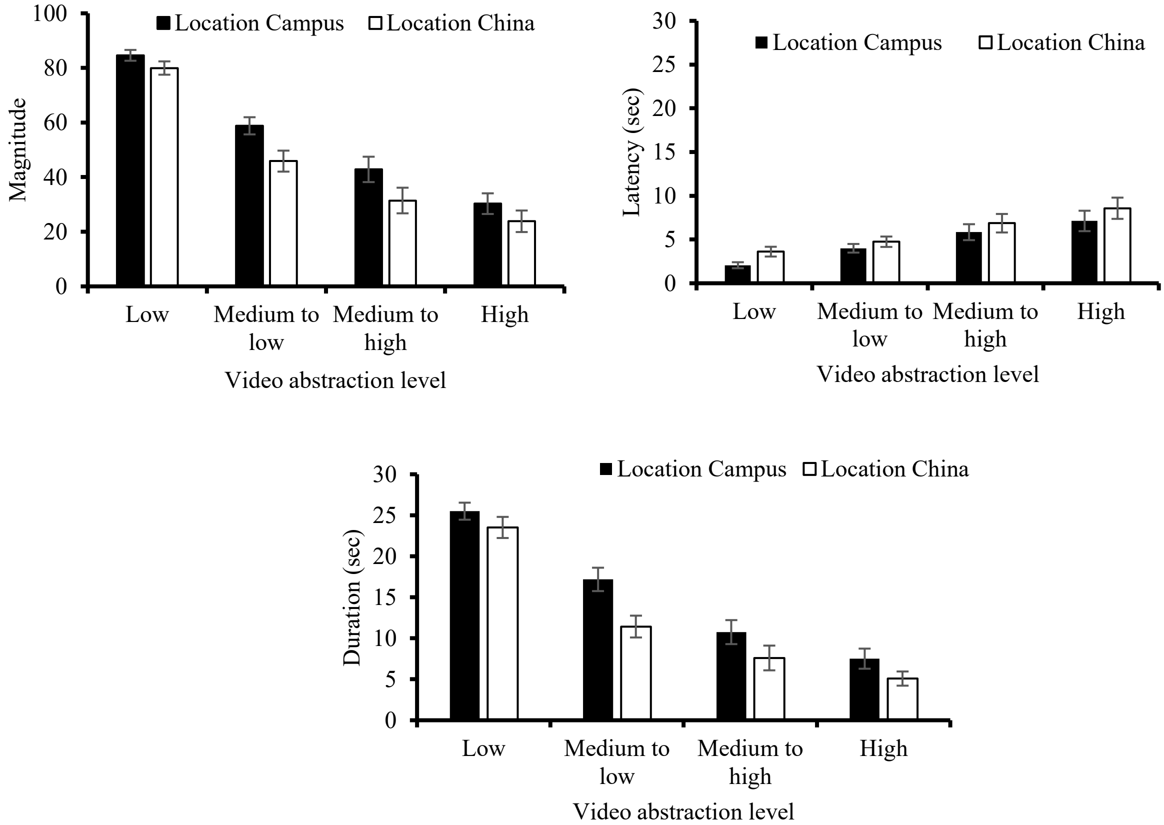


Fig. 3. Average vection magnitude, vection onset latency, and vection duration as a function of the different degrees of abstraction videos.

We conducted 4 (video abstraction) x 2 (video location) repeated-measures ANOVAs on the mean vection magnitude, latency and duration data (followed by post-hoc multiple comparisons). Significant main effects of video abstraction were found for vection magnitude (F (1.647, 26.348) = 130.342, $p < 0.001$, $\eta_p^2 = 0.891$), latency (F (1.587, 25.391) = 14.043, $p < 0.001$, $\eta_p^2 = 0.467$) and duration (F (3, 48) = 102.299, $p < 0.001$, $\eta_p^2 = 0.865$). Significant main effects of video location were also found for vection magnitude (F (1, 16) = 56.577, $p < 0.001$, $\eta_p^2 = 0.780$), latency (F (1, 16) = 7.298, $p = 0.016$, $\eta_p^2 = 0.313$) and duration (F (1, 16) = 58.798, $p < 0.001$, $\eta_p^2 = 0.786$). We also found that significant interactions between video abstraction and video location for vection magnitude (F (3, 48) = 3.321, $p = 0.027$, $\eta_p^2 = 0.172$) and duration

(F (3, 48) = 4.248, p = 0.01, η_p^2 = 0.210), but not for latency (F (3, 48) = 0.328, p = 0.805, η_p^2 = 0.02). Multiple comparisons revealed that there was no significant difference in vection duration between "medium to high" and "high" levels of video abstraction for China location stimuli, but such differences did reach significance for the Kyushu campus stimuli.

In this experiment, vection magnitude became weaker as the level of video abstraction increased (i.e., as it became more unnatural), which is generally consistent with previous findings of Schulte-Pelkum et al. (2003) and Riecke et al. (2006). However, we failed to find significant differences between "middle to low" and "middle to high", and between "middle to high" and "high" video abstraction conditions for vection latency – which suggests that the level of abstraction has little influence on the timing of vection onset. This finding is also consistent with those of Riecke et al. (2006) – who found that vection latencies for original/unmodified stimuli were shorter than that those for scrambled version of these stimuli. Importantly, vection onset latencies were shorter than 10 seconds for all of the conditions tested, which indicates that vection induction was not impaired by our CG technology based video manipulations (Sato et al., 2020). As was noted above, increasing the level of abstraction did however decrease the magnitude of vection intensity. Such findings could be related to the stimulus density. Previous studies have found that vection intensity decreases with the stimulus density (Keshavarz et al., 2019).

Relationships between naturalness, familiarity and vection intensity

Figure 4 below shows average ratings of naturalness and familiarity as a function of the level of video abstraction in each of video location conditions. Subjective estimates of naturalness and familiarity both decreased as the level of video abstraction increased. As expected, our participants were also less familiar with the modified stimuli generated from content filmed in China (compared to those generated from content filmed on their own local Kyushu campus).

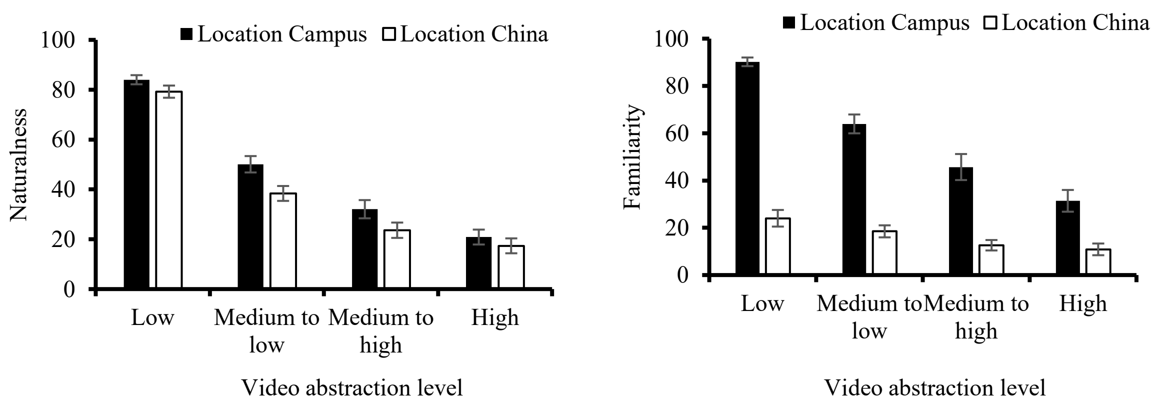


Fig. 4. Changes in the naturalness and familiarity of video stimuli as a function of the different degrees of abstraction videos.

We next conducted 4 (video abstraction) x 2 (video location) repeated-measures ANOVAs on the mean naturalness and familiarity rating data (followed by post-hoc multiple comparisons). We found significant interactions between video abstraction and video location for familiarity ratings (F (3, 48) = 43.833, $p < 0.001$, η_p^2 = 0.733), but not for naturalness ratings (F (1.854, 29.668) = 2.740, p = 0.084, η_p^2 = 0.146). According to Figure 4, the changes in vection intensity could have been related to subjective naturalness and familiarity. So, we conducted a correlation analysis which revealed that naturalness correlated significantly with all three vection indices (magnitude: r = 0.996, $p < 0.001$; latency: r = -0.929, p = 0.001; duration: r = 0.994, $p < 0.001$); We also found that familiarity correlated significantly with all three vection indices (magnitude: $r = 0.661, p = 0.074$; latency: $r = -0.728, p = 0.041$; duration: $r = 0.681, p = 0.063$).

As expected, we found that as the degree of abstraction increases, subjective ratings of naturalness decreases. However, subjective familiarity was also affected by the video location. As expected, the more natural and familiar the effect video stimuli, the stronger the vection intensity. Importantly, correlational analyses found that naturalness and familiarity both correlated significantly with all three of our indices of vection intensity. Moreover, we found that compared with familiarity, subjective naturalness had a higher correlation with vection intensity, which could better predict the change of vection intensity. This is consistent with previous research, which shows that vection intensity is affected not only by the physical properties of the stimulus but also by its subjective perceptual properties (such as visibility, brightness, perceived speed, perceived rigidity and so on – see Guo et al., 2021; Nakamura, 2019).

In summary, this study investigated the effects of different levels of video abstraction (generated by CG technology), and the effects of video familiarity, on vection strength. The results show that both of these factors affect vection intensity, especially the degree of abstraction. This may have something to do with subjective naturalness of the inducing stimuli. Thus, this study provides new evidence about the relationship between CG technology, video abstraction level and vection intensity.

Acknowledgements

This work was supported by JST SPRING, Grant Number JPMJSP2136. And the present study was partially supported by the "Diversity and Super Global Training Program for Female and Young Faculty (SENTAN-Q)", Kyushu University from MEXT.

References

Bossard, M., Goulon, C., & Mestre, D. (2020). Viewpoint oscillation improves the perception of distance travelled in static observers but not during treadmill walking. *Experimental Brain Research, 238*, 1–11.

Brandt, T., Dichgans, J., & Koenig, E. (1973). Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Experimental Brain Research, 16*(5),476-491.

Guo, X.R., Nakamura, S., Fujii, Y., Seno, T., & Palmisano, S. (2021). Effects of luminance contrast, averaged luminance and spatial frequency on vection. *Experimental Brain Research, 239*, 3507–3525.

Keshavarz, B., Philipp-Muller, A. E., Hemmerich, W., Riecke, B. E., & Campos, J. L. (2019). The effect of visual motion stimulus characteristics on vection and visually induced motion sickness. *Displays, 58*, 71-81.

Morimoto, Y., Sato, H., Hiramatsu, C. Seno, T. (2019). Material surface properties modulate vection strength. *Experimental Brain Research, 237*, 2675–2690.

Nakamura, S. (2019). Perceived rigidity significantly affects visually induced self-motion perception (vection). *Perception 48*(5):386–401.

Palmisano, S., Allison, R. S., Schira, M. M., & Barry, R. J. (2015). Future challenges for vection research: definitions, functional significance, measures, and neural bases. *Frontiers in psychology, 6*, 193.

Riecke, B. E., Schulte-Pelkum, J., Avraamides, M. N., Heyde, M. V. D., & Bülthoff, H. H. (2006). Cognitive Factors can Influence Self-Motion Perception (Vection) in Virtual Reality. *ACM Transactions on Applied Perception, 3*(3), 194-216.

Sato, H., Morimoto, Y., Remijn, G. B., & Seno, T. (2020). Differences in three vection indices (latency, duration and magnitude) induced by "camera-moving" and "object-moving" in a virtual computer graphics world, despite similarity in the retinal images. *i-Perception, 11*(5), 1–16.

Schulte-Pelkum, J., Riecke, B. E., von der Heyde, M., & Bülthoff, H. H. (2003). Circular vection is facilitated by a consistent photorealistic scene. *Proceedings of Presence 2003*.

Seno, T., Tokunaga, K., & Urasoko, R. (2018). Interviewing the professionals of making visual images on vection. *Vision, 30*(3), 95–102.

Seno, T., Sawai, K.I., Kanaya, H., Wakebe, T., Ogawa, M., Fujii, Y., Palmisano, S. (2017). The oscillating potential model of visually induced vection. *i-Perception, 8*(6).

Španel, M., & Kršek, P. (2006). Vector-based medical image segmentation using adaptive Delaunay triangulation. *In Proceedings of the Sixth IASTED International Conference on Visualization, Imaging, and Image Procesing*, Palma de Mallorca, ACTA Press.

Su, T., Wang, W., Lv, Z., Wu, W., & Li, X. (2016). Rapid Delaunay triangulation for randomly distributed point cloud data using adaptive Hilbert curve. *Computers & Graphics, 54*, 65-74.

Tokunaga, K., Ogawa, M., Ikehata, S., Masuda, T., Seno, T. (2016). Constitution of a database of vection scenes in Japanese movies and animations and experimental assessments of them. *Transactions of the Virtual Reality Society of Japan 21*: 35–47.

THE ROLE OF SPATIAL LOCATION IN THE IRRELEVANT SPEECH EFFECT REVISITED

Mitra Hassanzadeh, Wolfgang Ellermeier, Florian Kattner, Robyn Theuergarten, Marvin Heidinger, and Angelika Kern

Technical University of Darmstadt, Institute of Psychology, Darmstadt, Germany

<*mitra.hassanzadeh@tu-darmstadt.de*>

Abstract

The irrelevant speech effect (ISE) refers to the disruption of short-term memory by background speech that significantly impairs the recall of visually or auditorily presented content, such as lists of letters or numbers. While typically, this effect has been studied by presenting the interfering speech via headphones, with no, or at best primitive spatial cues (e.g., Jones & Macken, 1995), the present experiment used a "virtual reality" (VR) setup to position interfering speakers (both visibly and audibly) in space. In a "virtual café " scenario, participants listened to a target speaker while trying to ignore three distractor speakers positioned at different angles in space and producing either steady-state (repeating a single letter) or changing-state interference (a sequence of different letters). Preliminary results of an initial experiment in VR using both a fast (N=39 participants) and a slow rate of presentation (N=28 participants) for the interfering sounds suggest that while the classical 'changing-state' effect is obtained in VR as well, the expected 'streaming by location', i.e., participants exploiting the spatial layout to perceptually separate the interfering utterances by assigning them to different speakers is weak at best. Moreover, the prediction that the streaming effect should be greatest for the fast presentation rate is not born out by the present results.

EXPLAINING THE ANOMALOUS RESULTS OF STIMULUS COMPARISON: NOISY TARGET CORRELATES HELP OPTIMIZE DISCRIMINATION

Åke Hellström
Stockholm University
<hellst@psychology.su.se>

Abstract

The basic idea of weighting in stimulus comparison has become widely accepted. However, Hellström's (1979) sensation-weighting model (SWM) for stimulus comparison is still controversial. The full range of results of the experiment that the SWM was originally based on has remained neglected. Some of these results are "anomalous" and must be seen as a serious challenge for modelers who seek to apply currently popular ideas, such as Bayesian optimization, to stimulus comparison. Here, based on the SWM, the weighing-in of reference levels for discrimination optimization is discussed, emphasizing the potential role of noisy target correlates (NTCs).

In the comparison of successive stimuli, time-order errors (TOEs), occur, where the first (positive TOE) or the second (negative TOE) stimulus is systematically overestimated in relation to the other (Fechner, 1860). The TOE has often been found to change from positive to negative with increased magnitude level. This result has been explained in terms of the internal representation of the first stimulus being drawn toward the center of the stimulus range. In recent days, this notion has been framed in terms of Bayesian optimal estimation of the true magnitude of the first stimulus. But, as we shall see, the empirical base of this is shaky.

Reference Levels and Sensation Weighting

The thinking of Helson (Michels & Helson, 1954) was focused on adaptation to current stimulation in the form of the adaptation level (AL), a pooled average of past stimulation. This would optimize the measuring range. In stimulus comparison, the first stimulus would be weighted together with the AL, while the second stimulus is perceived directly. More recent views have likewise invoked the general notion of weighting-in of extra information in the form of some kind of reference level (ReL). In currently popular thinking (e.g., Raviv et al., 2012) the AL has been replaced as a ReL by the *Bayesian prior*, the weighting-in of which allows optimal estimation of the true magnitude of the first stimulus.

In Hellström (1979), successive three-category loudness comparisons were made for 16 loudness combinations under 16 temporal conditions. The mean and SD of the subjective difference *D* was estimated for each loudness combination. One main feature of the results was that with brief stimulus durations and interstimulus intervals (ISIs) the relation between stimulus magnitude level and the TOE was positive instead of negative (see Fig. 1) – an anomalous, or rather anti-conventional, result. Similar results were found by Hellström (2003).

Hellström (1979) described his results by the sensation weighting model (SWM),

$$d_{12}^{*} = u\,[s_1 \cdot \psi_1 + (1 - s_1)\,\psi_{r1}] - [s_2 \cdot \psi_2 + (1 - s_2)\,\psi_{r2}] + b, \qquad (1)$$

where d_{12}^* is the subjective difference between the first and the second stimulus, s_1 and s_2 are the weighting coefficients of the stimuli, and ψ_{r1} and ψ_{r2} are their current ReLs. Possible judgment bias is represented by b, and u is a scale constant.

The SWM was fitted to the results of Hellström (1979) with excellent fit, yielding weights for the first and second stimulus that varied greatly across conditions. Bayesian estimation of the first stimulus magnitude, as well as Michels and Helson's (1954) model, would predict, for all conditions, $s_1 < 1$ and $s_2 \approx 1$. Instead, with decreasing ISI, s_1/s_2 changed from < 1 to > 1, and the size of both weights decreased below 1 (see Fig. 1).

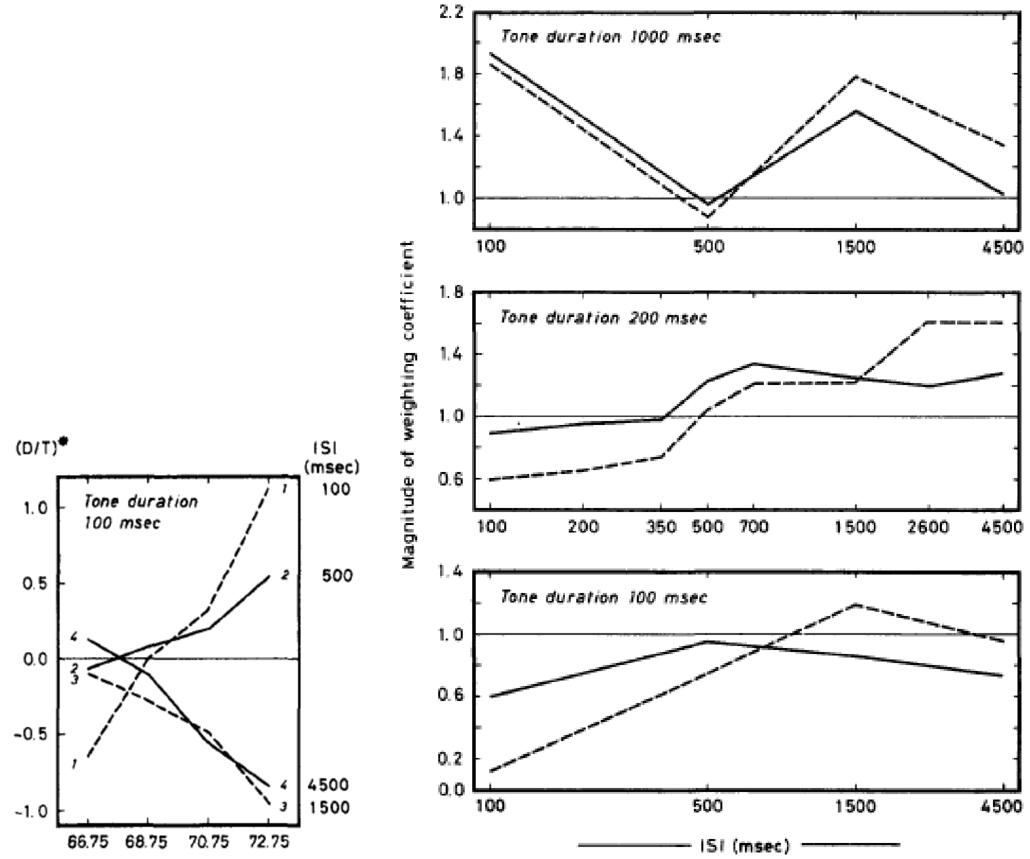


Fig. 1. Left: Mean subjective loudness difference versus mean tone loudness (only results for tone duration of 100 ms are shown). Right: SWM weight magnitudes s_1 (continuous) and s_2 (dashed), fitted to the loudness comparison data. From Hellström (1979).

Also, the SD of d_{12} varied greatly, being highest for conditions with large weight differences $|s_1 - s_2|$. Demonstrating the plausibility of the SWM, the results agreed well with the hypothesis that the measured d_{12} variances were actually variances of sensory magnitude differences weighted as per Equation 1, provided mean stimulus variances were roughly constant across conditions and the intrapair correlation was near 1. The latter result is reasonable, considering that the intrapair magnitude correlation due to the design was 0.88 in all conditions. The weighting, which varied so greatly across conditions, likely emerged to make the best possible, in terms of discriminability, out of the values of the other parameters – variances and intercorrelations of stimuli and ReLs. Further corroborating the SWM, in

Experiment 3 of Hellström et al. (2020) the model was put to a successful test by the manipulation of ReLs.

What is optimized in stimulus comparison?

To repeat and emphasize, results such as those of Hellström (1979) depicted in Fig. 1 seriously challenge such models (e.g., Raviv et al., 2012) that rest on the notion of Bayesian inference of the true magnitude of the first stimulus from its internal representation, which inevitably yields $s_1/s_2 < 1$. This problem is shared with the theories of Michels and Helson (1954) and Dyjas et al. (2012). These theories are useful as general accounts only if one neglects a great part of the empirical results (e.g., Hellström, 1979). The anomalous results tend to be found in conditions that are rarely used in stimulus comparison studies, perhaps due to their inconvenience (such as brief stimuli in rapid succession). To understand the full range of results, one must assume that both stimuli are in memory at the time of comparison (Fechner, 1860) and are affected by analogous processes (Hellström, 1979). However, this is not enough.

What is the core objective of quantitative perception? Centering and accurate estimation of magnitude are insufficient. Still more important is to stabilize perception and thereby maximize its sensitivity to magnitude changes that carry information – the signal-to-noise ratio (SNR) (Patching et al., 2012). This is done by minimizing the impact of noise. To achieve this, information from a noiseless ReL is of no help; instead the noise of the ReL must be correlated with that of the target. Such a ReL may be called a *noisy target correlate* (NTC).

Technological analogues of this mechanism are found in adaptive telescope optics as well as in noise-reducing sound-reproduction devices. In both cases, the principle is to weigh in an NTC. Using the ambient sound as an NTC for an auditive target, its noise is positively correlated with that in the target, so that subtracting, or weighing in negatively, this NTC will decrease the noisiness of the target. This also produces a contrast effect between target and NTC. One source of negative target-NTC correlation could be the vacillation of attention between a target and a ReL (Hellström, 1991), entailing out-of phase fluctuation of subjective magnitude of target and SNR (e.g., for subjective duration, loudness). If an NTC with negatively correlated noise is found, weighing it in with a positive weight reduces target noise, increasing the SNR. As a by-product, the target approaches, that is, assimilates to, the NTC. The contrast or assimilation effect shows up in the form of a systematic "error" such as a TOE.

For the perception of a single target stimulus, the weighting of target and reference for maximal SNR is described by a simple equation (Hellström, 1991). A more complicated equation for the optimization of SWM weights for maximal SNR in stimulus comparison was also developed (Hellström, 1986, 1989) and is described in the Appendix of Patching et al. (2012). The gain in SNR from using optimal weights can be huge. Simulations with the optimization equation show that a slight improvement in SNR is achieved by weighting the stimuli in roughly inverse relation to their SDs. With no NTC information, the weight sum is fixed to 2, but large changes in weight size, as empirically found, as well as big gains in SNR arise from weighting-in available NC information. Negative and positive NTC-target correlations yield s values in the SWM of < 1 and > 1, respectively.

The absolute size of the weights, which shifts with the ISI and stimulus duration, in particular for the second stimulus (see Figure 1) may be interpreted (e.g., Hellström, 1989) as indicating the quality of sensory processing. Likely, in the case of two brief stimuli with a short ISI, the processing of the two stimuli may interfere, causing low and fluctuating attention. This, in turn, may increase the variance of subjective magnitude and create negative NTC-target correlations. When the perceptual system takes these correlations into account (perhaps by heuristics or proxies) the s values are brought down in order to optimize discrimination.

Conclusion

To understand a "Proteus-like" (Fechner, 1860) perceptual phenomenon such as the TOE, it is necessary to take account of its full range of manifestations and not only those most conveniently elicited. Also, its explanation should be sought by considering the core goal of perception – to separate the message from the noise. This requires quite intricate and only partially known mechanisms, of which the TOE and related phenomena (Hellström, 1985) arise as by-products. Further investigation of these mechanisms is needed.

References

Dyjas, O., Bausenhart, K. M., & Ulrich, R. (2012). Trial-by-trial updating of an internal reference in discrimination tasks: Evidence from effects of stimulus order and trial sequence. *Attention, Perception, & Psychophysics, 74*, 1819–1841. doi: 10.3758/s13414-012-0362-4

Fechner, G. T. (1860). *Elemente der Psychophysik* [Elements of psychophysics]. Leipzig, Germany: Breitkopf & Härtel.

Hellström, Å. (1979). Time errors and differential sensation weighting. *Journal of Experimental Psychology: Human Perception and Performance, 5*(3), 460-477. doi: 10.1037/0096-1523.5.3.460

Hellström, Å. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin, 97*(1), 35-61. doi: 10.1037/0033-2909.97.1.35

Hellström, Å. (1986). Sensation weighting in comparing: A tool for optimizing discrimination. In Berglund, B., Berglund, U., & Teghtsoonian, R. (Eds.), *Fechner Day 86* (pp. 89-94). Stockholm: ISP.

Hellström, Å. (1989). What happens when we compare two successive stimuli? In G. Ljunggren & S. Dornic (Eds.), *Psychophysics in action* (pp. 25-39). Berlin: Springer.

Hellström, Å. (1991). Magnitude estimation of duration: Evidence of noise reduction by active perception. In *Fechner Day 91* (pp. 89-94). Durham, NC, USA.

Hellström, Å. (2003). Comparison is not just subtraction: Effects of time- and space-order on subjective stimulus difference. *Perception & Psychophysics, 65*(7)*, 1161-1177.* doi: 10.3758/BF03194842

Hellström, Å., Patching, G. R., & Rammsayer, T. H. (2020). Sensation weighting in duration discrimination: A univariate, multivariate, and varied-design study of presentation order effects. *Attention, Perception, & Psychophysics, 82*, 3196-3220. doi: 10.3758/s13414-020-01999-z

Michels, W. C., & Helson, H. (1954). A quantitative theory of time-order effects. *American Journal of Psychology, 67*, 327-334. doi: 10.2307/1418635

Patching, G. R., Englund, M. H., & Hellström, Å. (2012). Time-and space-order effects in timed discrimination of brightness and size of paired visual stimuli. *Journal of Experimental Psychology: Human Perception and Performance,* 36(4), 915–940. doi: 10.1037/a0027593

Raviv, O., Ahissar, M., & Loewenstein, Y. (2012). How recent history affects perception: The normative approach and its heuristic approximation. *PLoS Computational Biology, 8*(e1002731), 1–10. doi: 10.1371/journal.pcbi.1002731

AUDITORY STREAM SEGREGATION FOR COMPLEX TONES WITH SWITCHING FREQUENCY BANDS

Geng-Yan Jhang[1], Kazuo Ueda[2, 3], Hiroshige Takeichi[4], and Gerard B. Remijn[2]
[1]*Human Science International Course, Graduate School of Design, Kyushu University*
[2]*Department of Acoustic Design, Faculty of Design/Research Center for Applied Perceptual Science, Kyushu University*
[3]*Research and Development Center for Five-Sense Devices, Kyushu University*
[4]*Open Systems Information Science Team, Advanced Data Science Project (ADSP), RIKEN Information R&D and Strategy Headquarters (R-IH), RIKEN*
<jhang.geng-yan.266@s.kyushu-u.ac.jp, ueda@design.kyushu-u.ac.jp>

Abstract

Studies on auditory stream segregation have mainly focused on the proximity effects of frequency and onset-to-onset time. Here we introduced switching frequency bands of harmonic complex tones into the traditional experimental paradigm employing triplet tone sequences (Fig. 1), to examine how spectral switching interacts with frequency proximity between low (L) and high (H) tones in the range of 4 to 16 semitones. Experiment 1 extended the previous findings with pure tones, by shifting frequency of L tones from 1000 Hz to 200 Hz, shortening duration from 100 ms to 80 ms, and adding harmonic complex tones with 35 components. The results showed a trend of segregation for the pure-tone sequences similar to previous reports and more prominent segregation for the harmonic complex tones. Experiments 2 and 3 explored the interaction effects between spectral switching and intervals between L and H. Either even or odd numbered frequency bands in 2, 4, 8 or 16 frequency bands formed the L or H tones: odd-even-odd band sequences (Exp. 2) and even-odd-even band sequences (Exp. 3). The preliminary results (n = 3) showed that the 2-band conditions resulted in almost complete segregation irrespective of fundamental frequency proximity. Increasing the number of frequency bands tended to decrease the percentages of segregation for the L and H tones separated by the 4-semitone interval, and to a lesser extent by the 10-semitone interval, although some individual differences were observed. The results suggest that spectral switching interacts with frequency proximity on stream segregation, warranting further investigations.

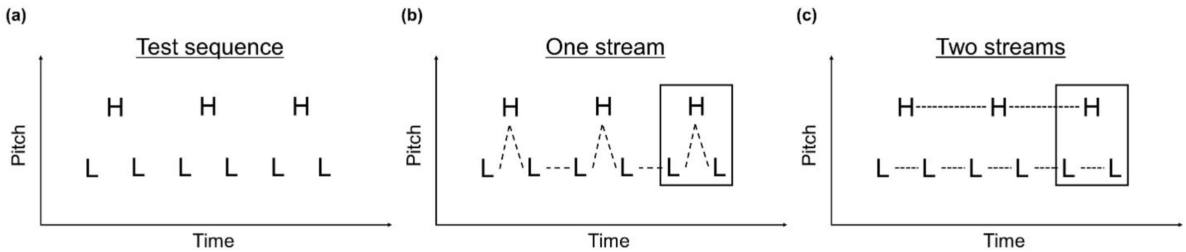


Fig. 1. Schematic illustrations of (a) a test sequence, (b) perception of one stream, i.e., integration, and (c) perception of two streams, i.e., segregation. Each "L" and "H" represents a low tone and a high tone. The dash lines represent perceptual connections between tones. The squares signify the final triplet (LHL) of the sequences that participants were instructed to report their percept.

A SPEEDED RECOGNITION MODEL OF INTENSITY CLASSIFICATION

Robert C. G. Johansson and Rolf Ulrich
Department of Psychology, University of Tübingen
<johansson@psycho.uni-tuebingen.de, ulrich@uni-tuebingen.de>

Abstract

A model describing how people categorize sensory events of different stimulus intensities is advanced. The central idea draws inspiration from mathematical conceptions of the sensory encoding stage according to which the output of peripheral stimulus transducers is a stream of afferent neural pulses traveling towards a central decisional loci. The bulk temporal properties of this neural time series is assumed to be broadly captured by the stationary Poisson process. We bridge the Poisson conjecture of sensory encoding with a sequential decision algorithm based on iterative likelihood ratio computation for each interpulse time recorded at the decision center. Overarching features of the theory are brought to bear upon empirical data gathered in a brightness classification task. It is demonstrated that a novel prediction of the model is a generalization of Pieron's law from simple reactions to the domain of binary choices.

Our ability to perceive and recognize physical stimuli of different intensities is a fundamental cognitive function. Yet, two independent research traditions seem to have developed roughly in parallel throughout the history of experimental psychology; a chronometric tradition concerned with the time-course of intensity processing, and a psychophysical tradition concerned with the acuity of intensity resolution in sensory pathways. In the chronometric tradition, stochastic processes unfolding in time have been invoked to account for various empirical benchmarks relating the speed of simple reactions to stimulus properties such as intensity, area and duration (Hildreth, 1979; Miller & Ulrich, 2003). The dependent variable of interest has naturally been response time (RT). The psychophysical tradition, in contrast, has been more concerned with the measurement of sensory thresholds and other mental indices derived from response probability. Signal Detection Theory (SDT; Green and Swets, 1966) and other statistical decision frameworks have often provided the analytical tools to investigate limits of stimulus detectability and discriminability.

A shortcoming of the above sketched division of research into separate traditions is an inability for current theories of intensity processing to account for both the speed and accuracy by which human observers tell apart stimuli of different intensity level (Link, 1992). The aim of the present paper is to present a novel theoretical framework which intuitively relates two-choice RT to the discriminability of a stimulus as a function of its intensity. More concretely, we are concerned with accounting for human performance in a type of perceptual recognition task henceforth termed *intensity classification*. In this task, the observer is faced with the following situation: on any given trial, he/she is presented with a physical stimulus of either weak or strong signal intensity. We term these two stimuli S_w and S_s, respectively. The observer is then to indicate as fast as possible which stimuli appeared by providing one out of two possible button presses.

Three challenges of immediate theoretical relevance seem to arise from this predicament: first, how is the physical intensity of the stimulus entered and represented in the nervous system? Second, what is the most efficient way for the observer to make use of the information contained within this sensory representation to distinguish the two signals? And finally, when has sufficient information been sampled so that the observer is ready to provide a manual response with some degree of confidence? The theoretical development presented below answers these three questions in turn by unifying two useful concepts from the literature. The first of these concepts is a general class of timing models of neural intensity coding as first advanced by Luce and Green (1972). When a stimulus event impends on the receptor surface of the body, peripheral sensory transducers produce a series of neural firings which travel along afferent fibers. Under the view of Luce & Green, the bulk temporal temporal properties of this neural time series is assumed to be well approximated by the stationary Poisson process. That is to say, the response characteristics of the sensory channel under consideration should adhere to Poisson's probability law

$$P[N(t) = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \tag{1}$$

where $P[N(t) = k]$ denotes the probability of recording k pulses during the observation interval t, and λ is a rate parameter representing the expected value and variance of the process. We shall in general assume that λ is monotonic with the intensity of the stimulus such that a high intensity signal S_s is associated with a faster rate of pulse generation λ_s and a weak signal S_w is associated with a slower rate of pulse generation $\lambda_w < \lambda_s$.

The second useful concept that we shall rely on in our theoretical development is the idea that human observers are intuitive statisticians, whose perceptual judgments reflect statistically optimal interpretations of sensory input under conditions of uncertainty. This notion is central to SDT, which uses statistical decision theory as the basis for distinguishing the influence of sensitivity and response bias on human detection of faint signals embedded in noise. Conveniently, it can be shown that the problem of distinguishing a signal from background noise is just a special case of recognizing one out of two prespecified signals (Tanner, 1956). As we shall see next, bridging the two domains of timing models and statistical decision theory leads to a particular case of the sequential likelihood ratio test (Wald, 1947) for a series of neural events under two mutually exclusive hypotheses.

The speeded recognition model

Following stimulus onset, the first pulse in the series is emitted, but the observer does not yet know the true identity of the stimulus, only that it is either of signals S_w or S_s. More technically, let the i-th pulse be recorded at time T_i ($i = 1, 2, \ldots$) following stimulus onset at $T_o = 0$. For each pulse there is a corresponding interpulse time $X_i = T_i - T_{i-1}$. Because the pulse train is generated by a latent Poisson process, it can be shown that X must itself be an exponentially distributed random variable with mean $\theta_w = 1/\lambda_w$ if the signal intensity is weak, or $\theta_s = 1/\lambda_s$ if the signal intensity is strong. We are then concerned with the chance of observing any interpulse time of duration x_i given a strong signal, divided over the chance of observing the same pulse time given a weak signal. Let Λ_i denote the log likelihood ratio for the evidence of a strong signal relative to the evidence of a weak signal as provided by the i-th interpulse time. Because the two exponential probability densities with rates θ_w and θ_s have the monotone likelihood ratio property in X, this quantity is given by

$$\begin{aligned}\Lambda_i &= -\log\left[\frac{L(|X = x_i)}{L(|X = x_i)}\right] \\ &= -\log\left[\frac{\frac{1}{\theta_s}\exp(-x_i/\theta_s)}{\frac{1}{\theta_w}\exp(-x_i/\theta_w)}\right]\end{aligned}$$

$$= \frac{\theta_s - \theta_w}{\theta_s \theta_w} \cdot x_i - \log\left(\frac{\theta_s}{\theta_w}\right). \tag{2}$$

The logarithmic operator assures that Λ_i is positive if x_i favors the strong signal hypothesis, and negative if it favors the weak signal hypothesis. In the special case where $\Lambda_i = 0$ the observation does not provide evidence for either hypothesis. This only happens at a single point where the two likelihood functions cross, i.e. where $L(\theta_s \mid X = x_i) = L(\theta_w \mid X = x_i)$.

A nice feature of working with the log likelihood ratio is that it is arithmetically convenient, because it simplifies the integration of information over time from a problem of products to a problem of sums. We shall assume the existence of a retention unit Σ which is able to hold this information as it accumulates over the course of an experimental trial. We restrict this evidence accumulation process to always have the same starting point at $\Sigma_0 = 0$. Following the arrival of a neural pulse at time T_i, the state of the retention unit is updated to

$$\Sigma_i = \Sigma_{i-1} + \Lambda_i. \tag{3}$$

It can be seen that Σ will enter a random walk over the real line that is discontinuous at each pulse time T_i. On average, we expect it to tend towards the positive domain if the stimulus is strong, and the negative domain if the stimulus is weak. But we do not want the observer to pool more and more information indefinitely without ever producing a behavioral response. We must therefore arrive at some reasonable stopping rule for interpulse time sampling. It has been shown by Wald (1947) that we can choose two barriers A and B a priori such that the type I and type II error rates of the sequential test procedure never exceed some desired constants of proportion α and β. Let

$$A(\alpha, \beta) \approx \frac{1-\beta}{\alpha} \tag{4}$$

and

$$B(\alpha, \beta) \approx \frac{\beta}{1-\alpha}. \tag{5}$$

We can now define a continue-sampling region such that the test procedure perpetuates whenever $A > \Sigma_i > B$. The observer chooses to terminate his sampling only if the sum of the log likelihoods overstep this region. When Σ_i exceeds A we say it is *absorbed* at A, and when it falls below B we say that the process is absorbed at B. Whenever absorption occurs at A, the observer responds that he/she saw a strong intensity target S_s. Correspondingly, a response indicating the presence of a weak target S_w is emitted when Σ_i exceeds B. Assume that the n-th pulse pushes the process towards absorption at either barrier A or B. The total decision time DT can then be expressed as

$$DT = \sum_{i=1}^{n} x_i \tag{6}$$

A conclusion of some importance can be reached from glancing at this equation, namely that the decision latency DT should in general increase in proportion to $E[X]$. This means that choice reaction time should be shorter for bright signals relative to dim signals when the absorption barriers are about equal in absolute magnitude. Analogous decreases in detection latency with increasing stimulus intensity has the status of psychophysical law in the literature on simple RT (Piéron's law). To our knowledge, this is the first time that this law has been extended to the domain of binary choices from first principles. Indeed, this prediction is such a general feature of the model that we would be in serious trouble if it was not to be confirmed empirically.

The final step necessary to yield sensible predictions of human performance is to account for those stages in the processing chain which are not strictly decisional in nature. These residual processes encompass transmission times in the nervous system, the motor time required for the observer to manually provide a response, and so forth. To this end, we add a residual time component t_r to *DT*. We make two debatable yet simplifying assumptions about t_r namely that 1) the component is strictly independent of the stimulus intensity of the target, and 2) its trial-to-trial variability is negligible. Hence

$$RT = DT + t_r \tag{7}$$

where t_r is an empirically estimated constant. We now turn our attention to an empirical test of the model. The aim of the experiment which will be described next was to generate data that would allow us to test the predictions of the speeded recognition model in terms of the entire constellation of classic empirical benchmarks; in particular, the shapes of the RT distributions belonging to correct and incorrect responses.

Methods

Participants

The first author R.J. (right handed, male, 29 years old) served as participant in the experiment.

Stimuli and apparatus

A Esprimo P956/E90+ computer (Fujitsu Limited, Tokyo) running a 32-bit Windows 7 operating system (Microsoft Corporation, 2009) controlled the experimental script in PsychoPy (Pierce & MacAskill, 2018). Visual stimuli were presented on a 24.1-inch FlexScan EV2495 LCD monitor (EIZO Corporation, Hakusan) with a pixel resolution of 1900 × 1200 and 60 Hz refresh rate. The stimuli were square patches of white light of either low (3 cd/m^2) or high (30 cd/m^2) intensity that measured 3 cm width × 3 cm height. The stimuli were presented centrally on the monitor for a duration of 300 milliseconds (msec). Background illumination was held constant at 0.3 cd/m^2 throughout the entire procedure. Luminous intensity of stimulus materials was measured with a P-9201-TF photometer (Gigahertz Optik, Türkenfeld). Response times were recorded from stimulus onset to the registering of the participants input via a custom set of response keys. A BlackBox Toolkit (Version 2) for external chronometry was used to verify the timing of stimulus presentation and response recording. The experiment was conducted in a sound- and light-attenuated booth with a chinrest fixed at 60 cm distance from the screen.

Procedure

Participant R.J. took part in 12 experimental sessions of about 45 min length each. Every session comprised 600 experimental trials divided over 6 experimental blocks. Before the start of each block, a set of 20 practice trials were presented. The mapping between stimulus intensity (dim or bright) and lateral response set (left or right button) was swapped between sessions. The start of a trial was signaled by the appearance of a faint fixation cross in the center of the screen for 500 msec. Following fixation offset, a target which was either dim or bright appeared on the screen after an additional 500 msec. The task was to determine the identity of the target with a speeded button press. The dependent measures were response time (time between stimulus onset and registering of the participants response) and response accuracy (correct or incorrect). In all respects, the experiment was conducted in accordance with the rules and regulations laid down by University of Tübingen's local ethics board and the European Research Council's ethics committee for human subject research.

Data Analysis

First, the data was inspected for anticipations (RTs < 200 msec) and attentional lapses (RTs > 2,000 msec). Then a set of 10 equispaced quantile values (0.05, 0.15, … , 0.95) were computed for each of four empirical RT distributions (dim or bright signal × correct or incorrect response). Finally, simulated data was fitted to empirical RT distributions in iterative fashion using the downhill simplex method (Nelder & Mead, 1965) until results were deemed satisfactory. Model fits are reported below in terms of the sum of the squared residuals between empirical and simulated quantiles, weighted by the inverse square root of the number of observations in each category according to the formula

$$SSR_w = \sum_{j=1}^{2} \frac{w_{S_j,r_1} \sum_{q=1}^{10} (O_q - E_q|S_j, r_1)^2 + w_{S_j,r_0} \sum_{q=1}^{10} (O_q - E_q|S_j, r_0)^2}{w_{S_j,r_1} + w_{S_j,r_0}} \quad (8)$$

where $w_{Sj,\, r1}$ is the weight for response option r_1 (r_1 =correct, r_0 =incorrect) to the j-th stimulus (dim or bright) and O_q and E_q are the observed and expected values of the q-th quantile point.

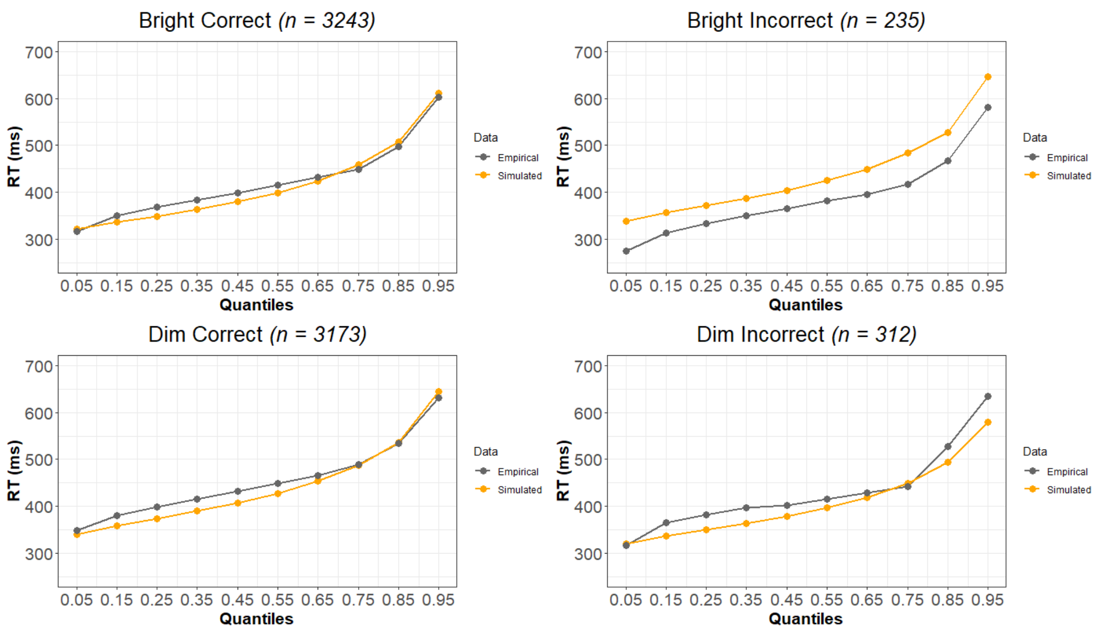


Fig 1. The plots show RT quantiles for correct and incorrect responses (left and right panels) to bright and dim targets (upper and lower panels). The number of empirical observations in each category is denoted *n*. Theoretical quantiles are computed from 20,000 simulated RTs.

Results and Discussion

Very few responses were too slow (0%) or too fast (less than 0.01%). Thresholds computed on the basis of error rates were A = 2.34 and B = -2.6. Figure 1 depicts empirical (gray) and simulated (yellow) quantiles belonging to all four RT distributions. The best solution afforded by the optimization procedure gave a model fit of SSR_w = 8581. This yielded parameter estimates $\theta_w = 1.22$, $\theta_s = 1.00$ and $t_r = 295$ reported here in msec. Empirical and simulated RT distributions are shown in Figure 1.

Three conclusions can be drawn from these results. First, the bright stimulus was classified faster (25-30 msec) than the dim stimulus: a pattern of performance that is well captured by the simulated data. We have therefore seemingly corroborated one of the models most immediately evident qualitative predictions, namely that choice RT should decrease as a function of stimulus intensity. Nonetheless, it remains to be determined whether parts of this intensity effect can be attributed to strictly attitudinal factors (response bias) because the data also indicates an increased propensity towards error on those trials where the target is dim (corresponding to a diminished separation between the bright absorption threshold A and starting point). Second, correct responses were in general slower than incorrect responses. It can intuitively be shown why the model predictions are in agreement with this finding if one considers how the standard error of the mean of a set of observations is inversely proportional to the number of observations in the sample. Samples of very few pulses are therefore more likely to lead the observer to hasty generalizations about the latent pulse-generating process, hence committing an error. The third conclusion is that the model predictions do not capture the ordering of incorrect response times; the model predicts faster incorrect responses when the stimulus is dim rather than bright, but we observe the opposite pattern of results. In our experience, this is a nonnegotiable feature of the model; it can partially accommodate either of the two incorrect response time distributions, but never both. The consequences of this misprediction can be clearly seen in Figure 1.

It bears noting that the estimated values of the parameters θ_w and θ_s imply neural oscillation rates of about 1,000 Hz. This figure is at odds with realistic firing rates of single units in the visual pathway by an order of magnitude. Yet it is our conclusion reached over the course of engaging with the model that relatively short interpulse times are necessitated, because it otherwise exhibits a systematic tendency to underpredict error rates. For firing rates of about 100 Hz the predicted error rates were 2-3% smaller than desired. For 1,000 Hz, the same figures are off by < 1%. One could therefore conceivably think of the estimated rate parameters as reflecting the superposed output of many receptive units transmitting in parallel.

In closing, we would like to comment on how the speeded recognition model relates to perceptual judgment more broadly. The purpose of this research was to shed light on aspects of human information processing pertinent to binary intensity classification. We therefore restricted our analysis to the case of statistically optimal frequency estimation of neural time series. Nevertheless, such a time series most certainly carries information about other stimulus attributes such as wavelength, size, spatial locus and so forth. Accordingly, the transmission speed of these attributes within the nervous system might similarly be constrained by pulse rate. Extensions of the present framework might therefore be construed such as to account for behavioral performance in other types of speeded recognition and detection tasks. We would however advise caution in such an enterprise. Generalizations of the model to arbitrary domains of sensation might inadvertently undermine parameter interpretability and obscure the mental processes subserving perceptual judgment.

Acknowledgements

The present research was funded as part of the DFG (Deutsche Forschungsgemeinschaft) funded research unit FOR2078 Modal and amodal cognition: Functions and interactions.

References

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.

Hildreth, J. D. (1979). Bloch's law and a poisson counting model for simple reaction time to light. *Perception & Psychophysics*, *26*, 153–162.

Link, S. W. (1992). *The wave theory of difference and similarity*. Lawrence Erlbaum Associates, Inc.

Luce, R. D., & Green, D. M. (1972). A neural timing theory for response times and the psychophysics of intensity. *Psychological Review*, *79*, 14–57.

Miller, J., & Ulrich, R. (2003). Simple reaction time and statistical facilitation: A parallel grains model. *Cognitive Psychology*, *46*, 101–151.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, *7*, 308–313.

Pierce, J., & MacAskill, M. (2018). *Building experiments in psychopy*. Sage Publishing.

Tanner, J., W. P. (1956). Theory of recognition. *Journal of the Acoustical Society of America*, *28*, 882–888.

Wald, A. (1947). *Sequential analysis*. John Wiley.

CONTINUOUS VERSUS GLOBAL JUDGMENT OF MUSICALLY INDUCED EMOTIONS

Lina Jost, Angelika Kern, and Wolfgang Ellermeier
Institute of Psychology, Applied Cognitive Psychology
Technical University Darmstadt, 64283 *Darmstadt, Germany*
<linaanastasia.jost@stud.tu-darmstadt.de>

Abstract

The beauty of music lies in its spectro-temporal structure stimulating our brain, therefore the instantaneous perception and the process of merging it into an overall judgment are of interest to get a deeper understanding of the processing of musically induced emotions. This study investigated the relationship between the continuous and the global judgment of felt emotions induced by four different pieces of film music which were rated on three dimensions: Valence, Arousal and Emotional Strength. Each subject listened to all musical stimuli in randomized order, first producing a continuous rating, one dimension at a time, by moving a slider on a bipolar scale with data being recorded at 1 Hz. Subsequently, a global rating was made on the same slider scale. Statistical data analysis ($N = 50$) showed a high positive correlation between the global judgment and the averaged continuous ratings for all three dimensions ($r_{rmValence}(148) = 0.92$; $r_{rmArousal}(149) = 0.93$; $r_{rmEmStrength}(149) = 0.79$). Thus, the overall impression of a musical piece appears to be an adequate representation of its moment-to-moment perception, and a better predictor than peaks or valleys in the elicited emotional qualities. Furthermore, the different musical pieces resulted in distinct emotional response profiles representing key musical features. Averaged curves for each of the three dimensions were largely parallel, but indicated that listeners discriminated between the attributes being judged. The results thus show that continuous tracking of emotional qualities constitutes a valid method for studying the elicitation and time course of music-induced emotions.

BINARY DECISIONS: MATHEMATICAL MODELS & APPLICATIONS

Diana Eugenie Kornbrot and George J. Georgiou
University of Hertfordshire
<d.e.kornbrot@herts.ac.uk>

Abstract

There has been substantial progress in creating and testing powerful models of binary decision making in many domains. But so far there is no consensus as to the most successful model in any specific perceptual domain. Furthermore, in spite of the potential there has been minimal research in potentially relevant applied domains such as medical decision making, evaluation of products and procedures and implicit bias (to name but a few). One reason is the lack of easy-to-use software to estimate parameter, and accepted standards if data deposited. This Ms. provides an overview of relevant parameters, an agenda for evaluating and promulgating theoretical, results, and summary of potential applications. It is very much a 'work in progress'

Models of binary decision making are a major achievement of mathematical psychophysical modelling. However, progress has been limited for three main reasons. Firstly, there is no comprehensive summary of theoretical findings according to domain (perceptual task, word recognition); or model type, (decision diffusion model (DDM), linear ballistic accumulator (LBA); or reaction time distribution fitting criteria (via moments, percentiles, or time grid). Secondly, there are no easy-to-use software packages, or standard methods of deposition of supplementary material (raw data, parameters estimates at an individual level, or code). Thirdly, these methods have rarely, if ever, been used for applied problems.

This Ms. addresses all three problems, starting with a general overview of information accrual models and their parameters; followed by potential areas of application and ending with a wish list of what is needed to solve problem 2 in terms of software and data deposition.

Information Accrual Models Structure and Parameters

The basic task is to assign one of two presented stimuli (S1, S2) to two alternative responses (R1, R2). R1 is correct for S1 and R2 is correct for S2.

All information accrual models are random walks and assume that people accumulate information at a specified drift rate (**v_i**) towards each of two possible response barriers, separated by a distance (**A)**. Figure 1. shows a general representation as implemented in the DStartM software (van den Bergh et al., 2020). The drift rate is **v_1** if stimulus 1 is presented and **v_2** if stimulus 2 is presented. Participants have no control over **v_1** and **v_2**, which depend on task difficulty (s), higher **v** for easier tasks or more talented people. Participants control time spent collecting information for each response, and the start point for information accrual, **z,** where **z = A/2** is 'unbiassedl"**.** Non-decision time (**T_{er}**). So measured **RT** = **T_{er} + d (decision time)**.

More sophisticated models include variability: with normally distributed drift (same sd **v_s** for both stimuli) and start point (sd = z_s) and **T_{er}** uniformly distributed with range **t_s.** More sophisticated methods and/or start points available with DstarM (van den Bergh et al., 2020).

None of the formulations I have seen follow earlier work showing sequential effects of previous trials (Laming, 1968). Preliminary analysis of motion detection (Dutilh et al., 2018) and word recognition (Wagenmakers, 2008) do show sequential effects with drift rate depending on previous trial accuracy (Kornbrot, 2021) Such models may obviate the need for *random* variation in drift rate (**v_s**). A matter for future research.

Model Structure

The diffusion decision model, DDM assumes $v_2 = -v_1$. The linear ballistic accumulator (LBA) makes no such assumption. I know of no global reviews comparing DDM with LBA

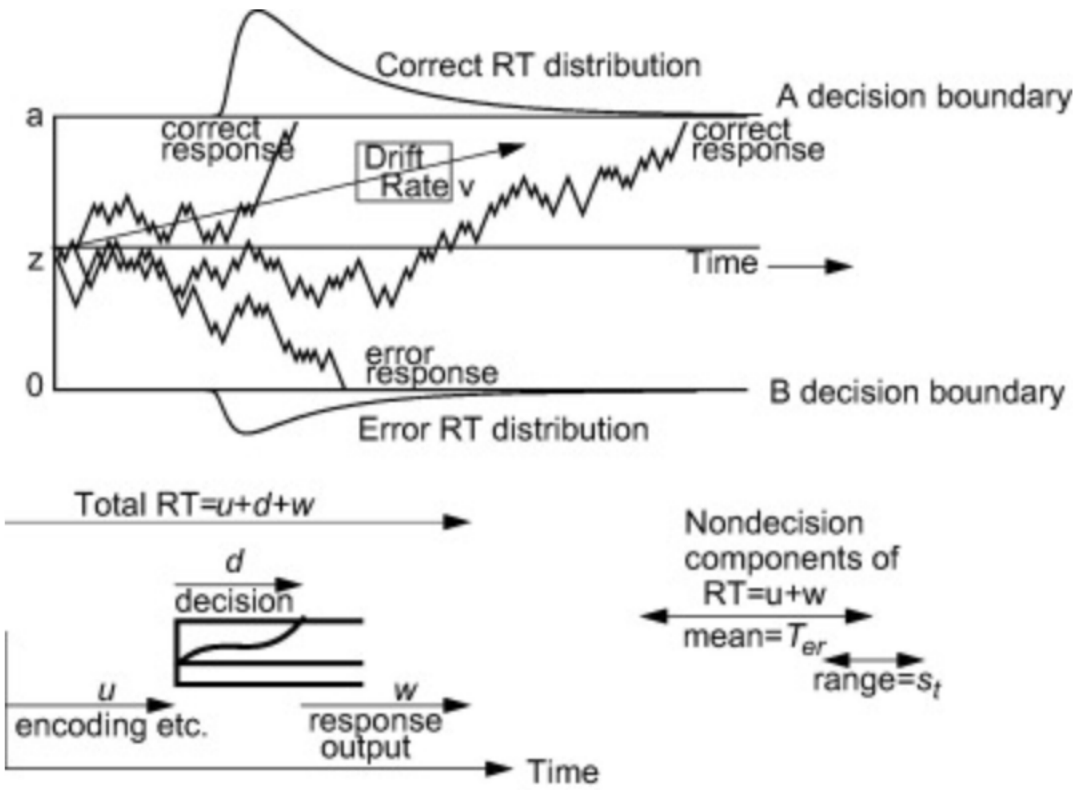


Fig. 1. General information accrual model from (Ratcliff, 2015)

Experimental Manipulations

All models predict a speed accuracy trade off. Conditions with speed pressures will be faster, but less accurate than those with accuracy pressure because speed pressure cause a reduction **in** A (the distance between decision boundaries?. This manipulation *always* works. All models also predict that easier conditions will be faster and more accurate than harder conditions because easier conditions have higher drift rates. Potentially there could be a measure incorporating both drift rate and barrier separation. We do not know of any such measure.

Data from Ratcliff (Ratcliff, 2015) on 2 numerosity judgment tasks in 2 caution conditions (speed, accuracy) are used to illustrate common findings.

Parameters in the numerosity study example (Ratcliff, 2015)

Data kindly provided by Prof. Ratcliff (he and his team are always prompt to supply information for studies conducted before data deposit became required. Kudos).

Stimuli were numbers of asterisks in task 1 and equivalent 2-digit numbers in task 2. There were two possible responses: below or above a target number. (In fact, either 6 or 8 different stimuli giving 6 or 8 drift rates, for simplicity average drift rates were used. There were 12 parameters for each of 21 participants: mean drift rate **(v),** barrier separation **(A)** and non-decision time ($\mathbf{T_{er}}$) in each of 4 conditions (2 tasks, 2 speed-accuracy instructions).

- Barrier separation **A** was slower in the accuracy condition, 348 msec, than the speed condition, 308 msec. Cohen's d = 2.16 (ginormous). A highly replicable finding.
- Barrier separation **A** was not significantly affected by task, as predicted by theory.
- Drift rate was significantly affected by task, asterisks = 342 msec, digits = 290 msec, Cohen's d = .77. As predicted the asterisk display was harder than the digit display.
- Drift rate was significantly affected by caution, speed = 291 msec, accuracy = 342 msec, Cohen's d = .64. A. Not unexpected, People sample less for an easy task, lower **A.**
- The non-decision time, $\mathbf{T_{er}}$, like the barrier separations was strongly affected by caution, $\mathbf{T_{er}}$ = 308 msec in speed condition, $\mathbf{T_{er}}$ = 348msec in accuracy condition. Cohen's d =1.5, a massive effect. Many of the analyses reported in the Dutilh (Dutilh et al., 2018) study showed an effect of caution on $\mathbf{T_{er}}$.

All 12 variables were entered into a factor analysis with intriguing results. Three factors accounted for 70% of variance (not bad). The first factor loaded on drift rates, v_i for all responses and tasks and also on barrier separation, A, for the asterisk task, A second factor had loadings from $\mathbf{T_{er}}$ in both task and barrier separation in the digit task.

Implications of experimental manipulation example for research

Most crucially, model descriptions and inference should be made on estimates of *all* parameters for *all* conditions for *all* participants. It is a testable hypothesis that people have a fixed drift rate at any given moment determined by the difficulty of the task. It is important to know whether any improvement in discrimination from a training regime actually increases the rate of information gain, or merely increases barrier separation. Similarly, it is useful to know whether people take advantage of an easier task by reducing the time spent accumulating information. Factor analysis of relations between parameters may also be a useful tool to explore relations between parameters.

There does not currently seem to be any systematic review of these questions**.** How people adjust their time allocation (barriers) to easier or harder tasks is an important psychological question both theoretically and practically for time pressured operatives in all walks of life.

The effect of caution on *non-decision* time is not easily explained in current models. A theoretical explanation is needed that may depend on task and motivation features.

Potential Applications

Binary decisions are widespread. E.g., screening medical images for disease and screening airports for terrorists. How does training affect performance? How about short breaks? Etc. (Beam et al., 1996; Sameti et al., 2009)

Another area of interest is 'implicit' bias, *time* to allocate diverse groups to atypical occupations or attitudes. *Accuracy* is ignored, effects may be due to shifts in barrier or drift rate (Greenwald, 20030728)

Tools and Communication

Supplementary material

In my view, progress will continue to be limited unless attention is paid to supplementary material that should be included with any information accrual model worth publishing.

Software

Currently, there are several R programs for information accrual models, often with options as to models being evaluated, the criteria for success, e.g. chi-square, Bayes, etc, and the distribution comparison (mean and variance (Wagenmakers, 2008), percentiles (Heathcote et al., 2002; Voss & Voss, 2007) or a time grid (Singmann, 2017; van den Bergh et al., 2020).

Software code deposit

What is required is code in a prevalent free language, such as R, with specified version and machine used; a list of required packages; the format of the data with column names that will be used in the R code as input parameters, the R code, a list of the output parameters generated by the code. An estimate of how long code takes to run. The reader of the Ms. should be able to replicate Ms. results using the code and hence apply to their own data.
Data deposit: in .csv or .xls or .xlsx, NOT only as .Rdata in the code
Raw data. All of it: participant, stimulus, response, reaction time, condition.
Parameter data. All parameters estimated for each participant in each condition.

Summary

Information accrual models have unfulfilled potential. Unanswered theoretical questions include best model, best estimation, role of previous errors and modelling of T_{er}. Us in applied areas is unexplored territory. Hopefully, this contribution will lead to progress in both theory and application by use of a clear guide of needed open supplementary material and a simple example of what can be achieved when parameters are fully deposited.

References

Beam, C. A., Layde, P. M., & Sullivan, D. C. (1996). Variability in the Interpretation of Screening Mammograms by US Radiologists: Findings From a National Sample. *Archives of Internal Medicine*, *156*(2), 209–213. https://doi.org/10.1001/archinte.1996.00440020119016

Dutilh, G., Annis, J., Brown, S. D., Cassey, P., Evans, N. J., Grasman, R. P. P. P., Hawkins, G. E., Heathcote, A., Holmes, W. R., Krypotos, A.-M., Kupitz, C. N., Leite, F. P., Lerche, V., Lin, Y.-S., Logan, G. D., Palmeri, T. J., Starns, J. J., Trueblood, J. S., van Maanen, L., … Donkin, C. (2018). The Quality of Response Time Data Inference: A Blinded, Collaborative Assessment of the Validity of Cognitive Models. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-017-1417-2

Greenwald, A. G. (20030728). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197. https://doi.org/10.1037/0022-3514.85.2.197

Heathcote, A., Brown, S., & Mewhort, D. J. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review*, *9*(2), 394–401. Scopus. https://doi.org/10.3758/BF03196299

Kornbrot, D. (2021). Evaluating Binary Decisions: Models, Predictors, and Collaboration. *Proceedings of the 37th Annual Meeting of the International Society for Psychophysics Fechner Day 2021*, 25–28. https://proceedings.fechnerday.com/index.php/proceedings/article/download/622/607

Laming, D. R. J. (1968). *Information theory of choice reaction times*. Academic Press.

Ratcliff, R. (2015). Modeling individual differences in response time and accuracy in numeracy. *Cognition*, *137*, 115–136. https://doi.org/10.1016/j.cognition.2014.12.004

Sameti, M., Ward, R. K., Morgan-Parkes, J., & Palcic, B. (2009). Image Feature Extraction in the Last Screening Mammograms Prior to Detection of Breast Cancer. *IEEE Journal of Selected Topics in Signal Processing*, *3*(1), 46–52. https://doi.org/10.1109/JSTSP.2008.2011163

Singmann, H. (2017, November 26). *Diffusion/Wiener Model Analysis with brms – Part I: Introduction and Estimation*. Henrik Singmann - Computational Psychology. http://singmann.org/wiener-model-analysis-with-brms-part-i/

van den Bergh, D., Tuerlinckx, F., & Verdonck, S. (2020). DstarM: An R package for analyzing two-choice reaction time data with the D∗M method. *Behavior Research Methods*, *52*(2), 521–543. https://doi.org/10.3758/s13428-019-01249-7

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*(4), 767–775. https://doi.org/10.3758/bf03192967

Wagenmakers, E.-J. der M. (2008). EZ does it! Extensions of the EZ-diffusion model. *Psychonomic Bulletin & Review*, *15*(6), 1229–1235. https://doi.org/10.3758/pbr.15.6.1229

DIRECTLY MEASURED STIMULUS DIFFERENCES PREDICTING CHOICE RESPONSES

Stephen W. Link
McMaster University
<link@mcmaster.ca>

Abstract

Psychology experiments often require a large number of experimental trials, resulting in a mound of data. Think of Fechner's weight comparison experiments. In Fechner's Elemente der Psychophysik (1860), six standard weights ranging from 300g to 3000g were judged against either of two heavier comparison weights. For the two-handed weight liftings, with 1024 trials in each of four conditions, the total number of trials was 6x4x1024 = 24,576. For comparisons of the same weights using one or the other lifting hand across eight 512 trial blocks the number of trials was 6x8x512=24,576. The total number of trials was 49,152. Imagine an experimental procedure that would reduce the needed number of trials to 720. This paper describes how to achieve that delightful number.

Directly measured stimulus differences (DMSD) Link (2022) is a new application of comparative judgment theory that requires an observer to both choose between two stimuli and create a measure of the size of the difference between the two stimuli. In the experiment described below the difference between the salaries of two jobs is created by the observer lengthening a horizontal line on a computer-controlled display screen to correspond to the observer's magnitude of the difference. An experiment on judging the difference between salaries for various jobs provides data to illustrate rather amazing results.

Twenty McMaster University undergraduate student volunteers, ranging from 21 to 25 years of age, earned $10 each to participate. Each student evidenced corrected 20-20 vision in each eye. If the student met this vision requirement, the student received an instruction sheet describing the experimental task. Instructions included a list of 20 occupational definitions from the United States Bureau of Labor Statistics. After reading the instructions the student discussed with the experimenter any questions about the meanings of the job definitions or the experiment itself. The 16 jobs used for experimental trials were Accountant, Bartender, Bus Driver, Chemist, Computer Programmer, Dentist, Engineer, Farmer, Flight Attendant, Lawyer, Physician, Police Officer, Psychologist, Secretary, Teacher, and Welder. Four jobs, Architect, Judge, Reporters, and Waiter/Waitress served as practice stimuli. Details of the trial events are found in Link (2022).

Results from the experiment appear in Table 1. These are averages across 20 students who each made a single choice and salary difference response to a unique job pairing. When the response made by the student to comparing job a vs b was recorded, the opposite response was also entered in the student's paired comparison matrix for the comparison of b to a. Also, the measure of salary difference for the unchosen b to a was entered as a negative value versus the value for the positive job choice. The same job pair was never presented to the same student in a reverse presentation, and a job was never presented against itself. Thus, each student filled a 16x16 matrix of paired comparisons presented in random order with a single judgment for 120 unique pairs of jobs. Students were balanced so that for one student receiving the comparison a against b another student received pair b vs a. The two data matrices, choice, and salary difference are each, 16 x 16 x 20, 16 x 16 jobs, for 20 students.

In Table 1 the rows correspond to the16 jobs. Column 2 provides the average across the 20 students of response probabilities for choosing which of 16 jobs had the larger salary across the paired comparisons. Column 3 provides the average of the lengths chosen to define the salary difference between the two jobs presented to each student.

Table 1. Average data for 20 students and 16 jobs showing the probability p of choosing a job as having a higher salary than all the other jobs, the value of Aθ determined from these probabilities, and L*, the average salary difference from the geometric mean salary difference for that job.

	P	Aθ	L*
Acct	0.62	0.490	0.23
Bart	0.125	-1.946	-0.81
BusD	0.119	-2.002	-0.86
Chem	0.581	0.327	0.23
Comp	0.663	0.677	0.34
Dent	0.848	1.719	0.86
Engi	0.739	1.041	0.50
Farm	0.269	-1.000	-0.62
Flight	0.263	-1.030	-0.51
Lawy	0.866	1.866	0.90
Phys	0.884	2.031	0.96
Poli	0.419	-0.327	-0.18
Psyc	0.656	0.646	0.31
Secr	0.217	-1.283	-0.67
Teac	0.445	-0.221	-0.18
Weld	0.286	-0.915	-0.50

What meaning can be attached to the marginal proportions, the $p_{.j}$ values, shown in Column 2 of Table 1? Consider the simplest case of unbiased responding. As Link (1975, 1978b, 1992, 2020, 2022) showed each response probability p_{ij} for choosing between job i and job j may be replaced by the parameters that it depends upon; $A\theta_{ij}$ where A is the Total Amount of accumulated comparative difference required for the response "Greater", and θ_{ij} is the logarithmic ratio of the mean values for the internal Poisson distributions for the two stimuli, λ_{cj} the Comparison j and λ_{si} for the Standard i. The unbiased response probability for choosing Job j to have a larger salary than Job i is

$$p_{ij} = \frac{1}{1 + e^{-A\theta_{ij}}}. \quad (1)$$

This is recognized as the "logistic equation" widely used to fit experimental data in many different scientific and medical fields. The derivation of the logistic equation as a consequence of a sequential process of accumulation was first made by Link (1978).

The logarithmic ratio of probabilities, pij and (1-pij) provides a value for the unknown exponent in equation (1),

$$A\theta_{ij} = \ln[p_{ij}/(1-p_{ij})]. \quad (2)$$

The average of any column of $A\theta_{ij}$ values in the 16 x 16 matrix is an average across standards for the fixed comparison in column j. This produces a marginal value, $A\theta_{.j}$ that corresponds to a marginal probability, $p_{.j}$, for choosing comparison S_j versus all stimuli including itself. The question of what this marginal probability represents is answered by computing the average Aθ for a fixed Comparison, Sj.

$$\begin{aligned} A\theta_{.j} &= \frac{1}{n}\sum_{i=1}^{n} A\theta_{ij} \\ &= \frac{1}{n}A(n\, ln(\lambda_j)) - \frac{1}{n}A\sum_{i=1}^{n} ln(\lambda_i) \\ &= A \ln\lambda_j - A\frac{1}{n}\ln(\lambda_1\lambda_2 \ldots \lambda_n) \\ &= A\, ln(\frac{\lambda_j}{\lambda *}) \end{aligned} \tag{3}$$

where $\lambda^* = (\lambda_1\lambda_2\ldots\lambda_n)^{(1/n)}$ is the geometric mean of all n = 16 comparison mean values. The geometric mean of the n values of λ, λ*, is Helson's (1964) Adaptation Level with respect to the internal stimulus magnitudes. Equation 3 is an extension of the original Fechner Law (1860) relating sensation to the comparison of two stimuli. Can Aθ.j be a measure of a sensation?

Using these parameters provides a marginal probability, $P_{.j}$, representing a choice between the comparison stimulus represented by the column and a hypothetical standard equal to the geometric mean of all stimuli. Specifically, the marginal probability for job j is the logistic probability,

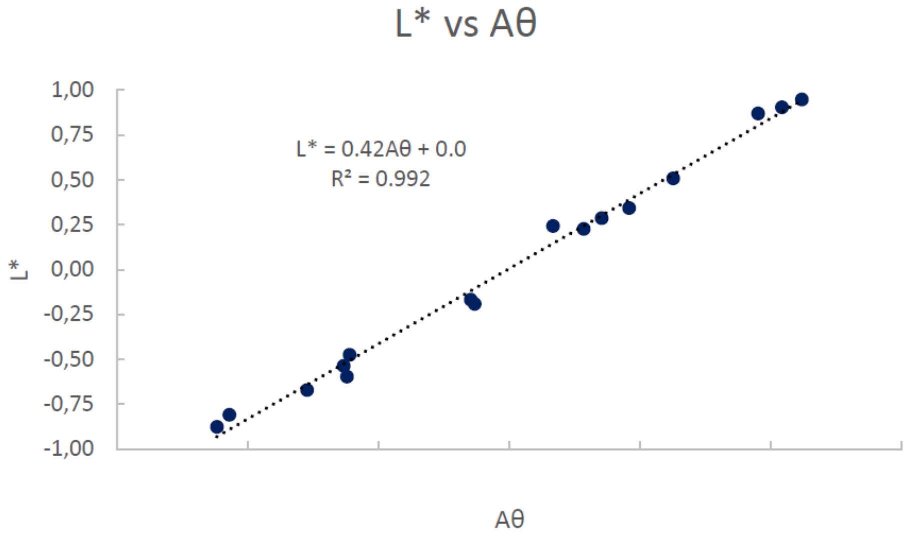
$$p_{.j} = \frac{1}{1 + e^{-A\theta_{.j}}}. \tag{4}$$

Therefore, the meaning of the marginal probabilities shown in Table 1 is that the magnitude of stimulus S_j is perceived to be of greater value than λ* with probability $p_{.j}$. Applying relation (1) to these values yields an estimate of $A\theta_{.j.}$ based on marginal probabilities,

$$A\theta_{.j} = \ln[p_{.j}/(1 - p_{.j})]. \tag{5}$$

The column means of the full Aθ paired comparison parameter matrix estimate values of internal magnitudes relative to the geometric mean intensity of all stimuli. How are these parameters related to stimulus differences? The value of $\theta_{ij} = \ln(\lambda_i/\lambda_i)$ may be thought of as both a ratio of stimulus magnitudes, or as the difference between logarithms of stimulus magnitudes. Furthermore, both quantities θ and A must enter the subject's evaluation of the existence of a difference between stimuli because the difference exists only after the subject decides a difference exists and makes a choice. By that point in time all the individual stimulus differences have vanished from the process of accumulating stimulus differences. All that is left is A, the value of the sum of all comparative differences and the subject's feeling of the size of a difference. To create a judgment both parameters are necessary but, at the end of the judgment process, what does the subject feel as a stimulus difference? A or θ or both?

Figure 1. The relation between Aθ and measures of length L* is based on the jobs in Table1.



The result:

$$L^* = k\,A\theta$$
$$= k\,A\ln(\lambda_c / \lambda_s)$$
$$= kA[\ln(\lambda_c) - \ln(\lambda_s)] \quad (6)$$

where k = 0.42, is an easily summarized law relating size of difference between two stimuli to the response threshold at A and the difference between logarithmic values for μ_s and μ_c. This, again, is the extension of Fechner's Law to the relation between the logarithm of two sensations multiplied by a constant. This is also a Power Law,

$$L^* = \ln(\lambda_c/\lambda_s)^{kA} \quad (7)$$

The constant, k, depends upon the student. There is much to be said about these relations.

Predicting individual choice behavior

To examine the ability of individual subject averages of DMSD to predict individual subject choices, a matrix of DMSD values was constructed for each individual subject. The DMSD entry in a particular cell measures the salary difference between two stimuli. The choice information is contained in the sign of DMSD. If positive, the job at the Top of a visual display was selected over the job at the Bottom of the visual display. Vice-versa, a negative value indicated a choice of the job at the Bottom rather than the Top of the visual display. Thus, even for a single choice, the signed value of DMSD defines both magnitude and direction of stimulus difference.

How well can these magnitudes of DMSDs predict individual choices? As a start, values of DMSD for each job were obtained by averaging values in each column of the individual paired comparison DMSD matrix. Each column contributes an average value characteristic of a particular job, as if compared against the geometric mean salary value for all jobs for a single student. Assuming these average DMSD values to measure the student's perception of the relative value of a job suggests that pairs of these average DMSD values be used to predict a student's individual choice. Theoretically, the student chooses the job with the greater DMSD.

More formally, let $L^*_{.jk}$ be the column average of DMSD values for job j for student k. The high correlation between L* and Aθ suggests that L* is a similarity transformation of Aθ. That is, there exists for each subject a scaling constant, w_k, such that

$$L^*_{.jk} = w_k A\theta_{.jk} .= w_k A\ln(\lambda_{.jk} / \lambda_{..k}^*) \quad (8)$$

where $\lambda_{..k}^*$ is the geometric mean of job values for student k.

A difference between two job L* values equals,

$$\begin{aligned} L^*_{.ik} - L^*_{.jk} &= w_k A\ln(\lambda_{.ik} / \lambda_{..k}^*) - w_k A\ln(\lambda_{.jk} / \lambda_{..k}^*) \\ &= w_k A\theta_{ijk} \\ &= L^*_{ijk} \end{aligned} \quad (9)$$

Theoretically, response probabilities for each cell of an individual paired comparison matrix may be predicted using Eq (8) and the distances L^*_{ijk} / w_k in place of $A\theta_{ijk}$.

When, as here, the linear relation between L* and Aθ is known, the constant w is the slope of a nearly perfect linear relation. These correlations are as large as 0.99 and as small as 0.88. Slopes vary from 1.78 to 3.01. The values of $L^*_{.jk}$ and w_k can now be used to predict individual student paired comparison choices.

Remember that for each job pair each of these subjects makes only a single binary choice, not the large number needed to estimate probabilities. Yet, it is still possible to predict individual subject choices, values of 0 if not chosen and 1 if chosen. Within the 16 x 16 choice matrix the best prediction of which of two jobs will be chosen is to select the job with the higher DMSD. For a single choice, this is the same as selecting the job with the higher marginal value of $L^*_{.jk}$. In particular, for jobs i and j, if $L^*_{.ik} > L^*_{.jk}$ then choose job i.

To ascertain the predictive value of these individual subject L* values average marginal DMSD values were compared pairwise with respect to size. The job with an average DMSD greater than another was predicted to be the subject's choice between the two jobs. In this way a predicted choice matrix of 240 job choices was computed for each of the 20 subjects. The proportion of correct predictions for each student ranges to 0.995 from a minimum of 0.935. Using the subjects' L* values yields a nearly perfect recovery of the original binary choice data. Additional results may be found in Link (2022). Luce (1959) would be gratified by this result.

Discussion

These many results suggest that directly measured stimulus differences provide a good measure of the subject's impression of stimuli. The method provides a new tool for the prediction of binary choice behavior. Applications of this measurement technique to other types of comparative judgments such as aesthetics, and its obvious extension to multidimensional scaling of DMSDs, provide for a deeper understanding of mental mechanisms and the stimulus representations on which much behavior is based.

The practical advantage of DMSDs over binary choices is that only a few are needed to establish relations between stimuli, and thus provide a basis for predicting binary choices. A complete paired comparison matrix requires at least n(n-1)/2 unique comparisons. Using DMSDs a single value within each row and column establishes a basis for prediction of choices for each cell in the entire matrix. A round robin of comparisons such as (Si, Si+1) (i=1, …, n; mod(n)) uses the immediate off-diagonal paired comparison matrix elements to provide single row and column estimates of marginal values. How many are needed to provide an accurate representation of the entire matrix is under investigation.

The fact that subject-produced line lengths can be used to predict response probabilities for the choice between two stimuli establishes a previously unsuspected link between the direct scaling methods of Stevens (1975), as illustrated by Parker and Schneider (1974), and the comparative difference ideas due to Fechner and Thurstone. The Poisson basis for paired comparisons using the Wave Theory ideas advanced by Link (1992, 2020, 2022), extended the sequential comparative difference model (Link 1975, Link and Heath, 1975), and advanced Fechner's amazing theory of subjective choice. Now, based on the Poisson based theory (1992), new derivations suggest using stimulus difference, as measured by subject-generated horizontal line lengths, for example, as the basis for choices between stimuli. The obtained correlation of 0.99 with 99% of the variance accounted for by the linear relation between estimates of process parameters and directly measured DMSD's opens the door for many new psychophysical investigations into our deeper mental processes. Direct measures of stimulus difference have a theoretical basis that permits prediction of individual choice behavior. Fechner and Stevens are united.

References

Fechner, G. T. (1860/1966). *Elemente der psychophysik*. (Trans. Adler, H. E., Howes, D. H., Boring, E. G.) Holt Rinehart Winston. NY.

Helson, H. (1966). *Adaptation-level theory*. NY: Harper and Row.

Link S. W. (1975). The relative judgment theory of two-choice response time. *Journal of Mathematical Psychology, 12*, 114-135

Link, S. W. (1978). *Relative judgment theory of the Psychometric Function*. In (J. Requin (Ed.) Attention and Performance VII (pp. 619-630). Hillsdale, NJ: Lawrence Erlbaum Associates.

Link, S. W. (1992/2020) *The wave theory of difference and similarity*. Hillsdale, NJ: Lawrence Erlbaum Associates/NY: Routledge.

Link, S. W. and Heath, R. H. (1975) A sequential theory of psychological discrimination. *Psychometrika, 40*, 77-105.

Luce. R. D. (1959). *Individual choice behavior*. NY: Wiley

Parker, S. and Schneider, B. (1974) Nonmetric scaling of loudness and pitch using similarity and difference measures. Perception and Psychophysics. V15, 2, 238-242.

Stevens, S. S. (1975). *Psychophysics: An introduction to its perceptual, neural and social prospects*. (G. Stevens, Ed.).NY: Wiley.

Thurstone, L. L. (1955). *The measurement of values*. Chicago: University of Chicago Press.

OLD AND NEW VIEWS ON RATIO JUDGMENT

Sergio Cesare Masin
University of Padova, Italy
<scm@unipd.it>

Abstract

Old and new psychophysical data suggest that judged sensory ratios fail to match the actual sensory ratios, and that this discrepancy is due to the incapacity of people to judge sensory ratios. Directions for future research are outlined.

To measure sensory magnitude, Merkel (1888) assumed that people are able to produce pairs of sensory magnitudes in the specific numerical ratio of 2:1, and Richardson (1929) more boldly assumed that people are able to estimate (that is, to express as a number) the ratio between any pair of sensory magnitudes. Stevens (1936) endorsed Richardson's assumption.

Productions and estimates of reciprocal sensory ratios

For each of different loudness levels of a standard sound *S*, Geiger and Firestone (1933) asked subjects to produce the loudness of a variable sound *V* that was, in turn, 1/2, 1/4, 1/10, 1/100, 2, 4, 10, and finally 100 times that of *S*. The authors exemplified their results by asserting that if subjects "are first asked to adjust the loudness of a sound to a value twice as loud, then are asked to adjust this sound to a value half as loud, they will not, on the average, arrive at the original sound, but to a value either louder than or less loud than the original sound" (p. 26). For any sensory magnitudes *a* and *b*, let J_1 and J_2 denote the estimates of the ratios a/b and b/a, respectively. These estimates are equivalent to the respective sensory ratios if $J_1 \cdot J_2 = 1$, given that $(a/b) \cdot (b/a) = 1$. Svenson & Åkesson (1966) found instead that $J_1 \cdot J_2 \neq 1$.

Test of Richardson's assumption

Engen and Levy (1955) and, perhaps surprisingly, Stevens (1956) himself found that estimation of sensory ratios yielded unreliable measures of sensory magnitude because the exponent of the psychophysical power function turned out to vary with the value of the standard stimulus used.

Arithmetical relationships between estimates of sensory ratios agree with this finding. For any sensory magnitudes *a*, *b*, and *c*, let J_1, J_2, and J_3 denote the estimates of the ratios a/b, a/c, and c/b, respectively. Because $a/b = (a/c) \cdot (c/b)$, these estimates yield reliable sensory measures only if $J_1 = J_2 \cdot J_3$. Eisler (1960) and Fagot and Stewart (1969) found instead that $J_1 \neq J_2 \cdot J_3$.

The Ross-Di Lollo hypothesis

The results of the above tests (and of a surplus of more recent similar tests not considered here) led to the belief that people misjudge sensory ratios (Luce, 2002; Shepard, 1978). Could this misjudgment indicate incapacity to produce or estimate sensory ratios? The empirical results of Ross and Di Lollo (1971) described below suggest a positive answer.

If subjects can estimate sensory ratios of variable stimuli *V* relative to a standard stimulus *S*, Richardson's assumption states that subjects produce the estimates

$$J = v/s \tag{1}$$

with v and s being the perceptual or memory magnitudes caused by V and S, respectively. If the measures of v and s are known, the functional relationship $J = v/s$ implies that plotting J in a Cartesian coordinate system against v separately for different values of s yields *factorial curves* diverging rightward. Unfortunately, v and s are unknown. However, considering that v increases monotonically with V, the relationship $J = v/s$ also implies that plotting J against V separately for each S yields factorial curves that diverge rightward.

With V and S known, a test of this implication serves as an indirect test of Richardson's assumption (Anderson, 1974). Since v and s are unknown, this test can only be an ordinal test: Richardson's assumption is rejected if the empirically-obtained factorial curves are parallel or converge rightward, and is supported—but not verified—if these curves diverge rightward. The test cannot verify Richardson's assumption because there are functional relationships other than $J = v/s$ that also imply a rightward divergence of factorial curves.

Ross and Di Lollo (1971) had three groups of subjects compare variable stimuli V relative to a single standard stimulus S. The range of V differed for each group, which involved a largely different remembered s for each group. Subjects were asked to estimate the ratio v/s. Plotting J against V separately for each s yielded nearly parallel factorial curves. Since estimates of $|v - s|$ imply parallel factorial curves, Ross and Di Lollo (1971) hypothesized that subjects manifested "interval scale behavior, that is, the direct representation of intervals as numbers by a process of comparison of intervals" (p. 520). That is, subjects did not estimate sensory ratios. Atkinson and Ward (1972), Fagot, Stewart, and Kleinknecht (1975), and Schneider, Parker, and Upenieks (1982) obtained data in line with this hypothesis.

Test of the Ross-Di Lollo hypothesis

On each trial of two experiments, Masin (2014) presented two successive stimuli: a surface of luminance V (8, 15, 26, or 49 cd/m^2) and a surface of luminance S (1, 2, or 4 cd/m^2) in one experiment, and an object of weight V (440, 535, 650, or 750 g) and an object of weight S (85, 175, or 260 g) in the other. For each combination of V and S, the subjects' task was to estimate the brightness ratio v/s in one experiment and the heaviness ratio v/s in the other. Instructions to make ratio estimates used no numerical examples to exclude spurious divergence of factorial curves (Guirao, 1987; Masin, Weiss, & Brancaccio, 2021; Robinson, 1976)

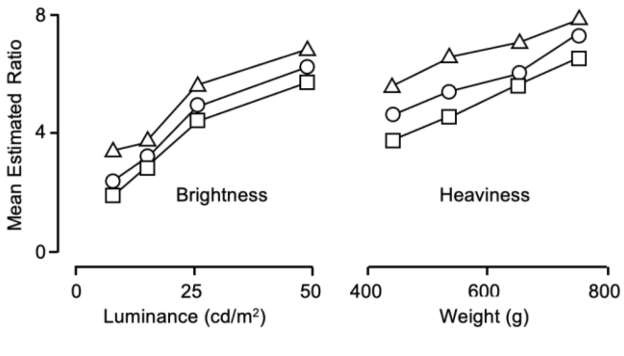


Fig. 1. Mean estimated brightness ratio or heaviness ratio as a function of luminance or weight, respectively. The curve parameter is the respective standard stimulus intensity. Triangles denote the smallest standard. Redrawn from data of Masin (2014, Figs. 1 and 2).

Fig. 1 shows the factorial curves obtained by plotting mean estimated brightness ratio and mean estimated heaviness ratio against the corresponding V, separately for each corresponding S. Triangles and squares denote the smallest and largest S, respectively. Factorial curves are nearly parallel. Because the relationship $J = |v - s|$ implies parallel factorial curves, the results agree with the Ross-Di Lollo hypothesis of "interval scale behavior." Ratio production yielded similar results (Masin, 2007, 2014).

Ratio judgment through counting

It may be that people easily believe that they can generally estimate sensory ratios because they often visually verify that they can somewhat accurately apprehend small ratios of visual extents. This automatic apprehension resembles subitizing, the ability to apprehend numerosities of 1 to 4 items at a glance (Taves, 1941; Balakrishnan & Ashby, 1991, 1992).

People subitize through an automatic counting process of which they are unaware. This counting process can, however, be demonstrated by measuring the mean time the subjects take to name the number of items presented in a visual display. The left diagram in Fig. 2 shows this mean response time plotted against item numerosity (Balakrishnan & Ashby, 1991). Mean response time increases with item numerosity indicating an automatic counting process.

People also use an automatic counting process to perform the task of estimating length ratios (Hartley, 1977). The central diagram in Fig. 2 shows the mean response time taken to name the apparent-length ratio between each of four variable lines of 2 to 5 cm and a standard line with mean length of 1 cm (Masin, 2013). Mean response time increases with the length of the variable line. This fact suggests that subjects automatically counted the number of times the standard line was contained in each variable line, rather than merely apprehending length ratios.

The Stevens-Moskowitz hypothesis

The following results indicate that the counting process used in the task of estimating length ratios might occur only for extensive sensations. The right diagram in Fig. 2 shows the mean time taken to name the brightness ratio between each of four variable luminances from 8 to 50 cd/m^2 and a standard mean luminance of 2.5 cd/m^2 (Masin, 2013). Mean response time was essentially invariant with luminance, indicating that subjects did not count brightness units. It might thus be that people cannot count units of intensive sensation.

If people fail to estimate brightness ratios (left diagram in Fig. 1) and to count brightness units (right diagram in Fig. 2), what other operation do they use to perform the task of estimating brightness ratios? Perhaps surprisingly, Stevens (1956) himself provided an answer, noting that

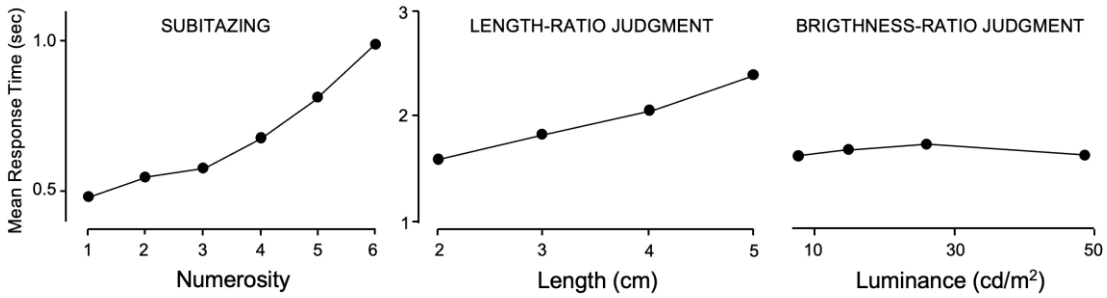


Fig. 2. Mean response time to report item numerosity (left) and to estimate length ratios (center) and brightness ratios (right) plotted against numerosity, length, and luminance, respectively. Redrawn from data of Balakrishnan and Ashby (1991, Fig. 1) and Masin (2013, Figs. 1 and 2).

"some [of the subjects instructed to estimate sensory ratios] seem to make their estimates on an interval-scale, or even on an ordinal scale, instead of on the ratio-scale we are trying to get them to use" (p. 23). Moskowitz (1977) also remarked that "[t]he range of numbers [ratio estimates] differs among individuals. Some people choose numbers between 1 and 10 and operate as if they were working on a category scale with equal intervals rather than equal ratios. Others choose a range between 0 and 100 and operate in the same way. Finally, others choose numbers between 0 and 1,000 (or some other range)" (p. 210).

Generality of the Stevens-Moskowitz hypothesis

The Stevens-Moskowitz hypothesis basically states that, if instructed to estimate sensory ratios, some subjects may instead rate sensory differences on a category scale. Masin (2022) tested the generality of this hypothesis, using instructions without numerical examples to prevent spurious effects possibly caused by such examples.

The experiment whose results are reported in the left diagram Fig. 1 was repeated using one range of *V* for one group of subjects, and about double this range for another group. If estimates of brightness ratios are proportional to brightness ratios, the height of factorial curves should increase with the range of *V*. If, instead, estimates of brightness ratios are ratings of brightness differences on a category scale, the height of factorial curves should be essentially the same for the two groups, considering that both groups would map brightness differences on scales with statistically the same endpoints. Fig. 3 illustrates the results. The patterns of factorial curves obtained by the two groups have the same height, indicating that subjects rated sensory differences on a category scale rather than estimating sensory ratios.

Different judgment operations

The results illustrated above suggest that people count units of apparent extent to fulfill the task of estimating ratios of extensive magnitudes (length) and that they rate differences of intensive magnitude to fulfill the task of estimating ratios of intensive magnitudes (brightness). It may be that people use various types of judgment operations (Masin, Brancaccio, & Tomassetti, 2019).

The results reported in the right diagram in Fig. 1 were obtained with subjects lifting pairs of weights unimanually. Masin and Brancaccio (2017) repeated this experiment with subjects now lifting pairs of weights bimanually. These pairs were factorial combinations of six variable weights *V* of 500 to 1000 g with four standard weights *S* of 100 to 400 g. Fig. 4 shows mean estimated heaviness ratio plotted against *V* separately for each *S*. The factorial curves diverge, thus indicating that subjects did not estimate heaviness differences. This divergence might seem

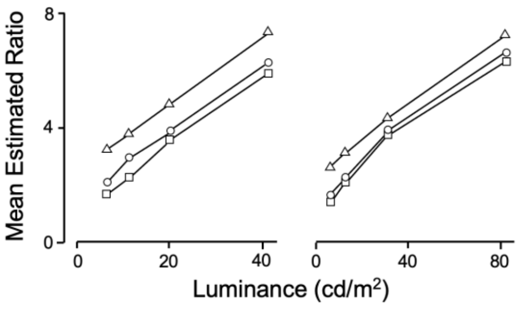


Fig. 3. Mean estimated brightness ratio plotted against variable stimulus luminance for each of three standard stimulus luminances (triangles denote the smallest standard) and for each of two different luminance ranges. Redrawn from data of Masin (2022, Fig. 2).

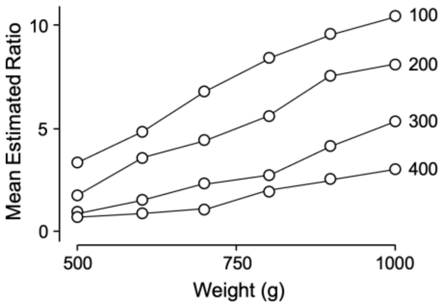


Fig. 4. Mean estimated heaviness ratio plotted against variable stimulus weight for each of four standard stimulus weights (grams). Redrawn from data of Masin and Brancaccio (2017, Fig. 3).

to support that subjects estimated heaviness ratios. However, the obtained factorial curves are essentially uniformly spaced, whereas the sensory relationships $J = v/s$ implies hyperbolically spaced factorial curves [the values used for S were equispaced and the psychophysical function for heaviness was essentially linear in the range of 100 to 400 g (J. C. Stevens & Rubin, 1970)]. These results thus indicate that subjects may have used a judgment operation based on a sensory relationship other than a sensory difference ($J = |v - s|$) or a sensory ratio ($J = v/s$).

Future research

The following may be relevant goals for future research: (i) determining which sensations allow for mental counting of sensory units, (ii) exploring comprehensively the sensory relationships people respond to while they fulfill the task of producing or estimating sensory ratios, and [since people can be conditioned to produce specific patterns of factorial curves (Price, Meyer, & Koh, 1992)] (iii) assessing new methods of constrained scaling where subjects learn to produce the patterns of factorial curves implied by the functional relationships $J = |v - s|$ or $J = v/s$.

References

Anderson, N. H. (1974). Algebraic models in perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2, pp. 215–298). New York: Academic Press.

Atkinson, W. H., & Ward, L. M. (1972). On the internal consistency of magnitude estimations. *Perceptual and Motor Skills*, 35, 903–913.

Balakrishnan, J. D., & Ashby, F. G. (1991). Is subitizing a unique numerical ability? *Perception & Psychophysics*, 50, 555–564.

Balakrishnan, J. D., & Ashby, F. G. (1992). Subitizing: Magical numbers or mere superstition?. *Psychological Research*, 54, 80–90.

Eisler, H. (1960). Similarity in the continuum of heaviness with some methodological and theoretical considerations. *Scandinavian Journal of Psychology*, 1, 69–81.

Engen, T., & Levy, N. (1955). The influence of standards on psychophysical judgments. *Perceptual and Motor Skills*, 5, 193–197.

Fagot, R. F., & Stewart, M. (1969). Tests of product and additive scaling axioms. *Perception & Psychophysics*, 5, 117–123.

Fagot, R. F., Stewart, M. R., & Kleinknecht, R. E. (1975). Representations for biased numerical judgments. *Perception & Psychophysics*, 17, 309–319.

Geiger, P. H., & Firestone, F. A. (1933). The estimation of fractional loudness. *Journal of the Acoustical Society of America*, 5, 25–30.

Guirao, M. (1987). Estimaciones numéricas de brillo y luminosidad con y sin rangos fijos. *Revista Latinoamericana de Psicología*, 19, 297–317.
Hartley, A. A. (1977). Mental measurement in the magnitude estimation of length. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 622–628.
Luce, R. D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review*, 109, 520–532.
Masin, S. C. (2007). The judgment operation of ratio production. *Perception & Psychophysics*, *69*, 1–8.
Masin, S. C. (2013). On the ability to directly evaluate sensory ratios. *Attention, Perception, & Psychophysics*, *75*, 194–204.
Masin, S. C. (2014). Test of the ratio judgment hypothesis. *Journal of General Psychology*, 141, 130–150.
Masin, S. C. (2022). Evidence that numerical estimates of subjective ratios may be numerical ratings of subjective differences. *Psicológica*, 43(1): e14568.
Masin, S. C., & Brancaccio, A. (2017). Judgments of differences and ratios of subjective heaviness. *Attention, Perception, & Psychophysics*, 79, 1860–1869.
Masin, S. C., Brancaccio, A., & Tomassetti, A. (2019). Tests of the abilities to judge ratios of extensive and intensive sensory magnitudes. *Attention, Perception, & Psychophysics*, 81, 2102–2119.
Masin, S. C., Weiss, D. J., Brancaccio, A. (2021). Further evidence of the invalidity of the magnitude estimation method. *European Review of Applied Psychology*, 71, 100649.
Merkel, J. (1888). Die Abhängigkeit zwischen Reiz und Empfindung, Erste Abtheilung. *Philosophische Studien*, 4, 541–594 [Zweite Abtheilung: 1889, 5, 245–291, 499–557].
Moskowitz, H. R. (1977). Magnitude estimation: Notes on what, how, when, and why to use it. *Journal of Food Quality*, 3, 195–227.
Price, P. C., Meyer, D. E., & Koh, K. (1992). Multivariable function learning: Applications of the adaptive regression model to intuitive physics. In J. K. Kruschke (Ed.), *Proceedings of the fourteenth annual conference of the Cognitive Science Society* (pp. 552–557). Hillsdale, NJ: Erlbaum.
Richardson, L. F. (1929). Quantitative mental estimates of light and colour. *British Journal of Psychology*, 20, 27–37.
Robinson, G. H. (1976). Biasing power law exponents by magnitude estimation instructions. *Perception & Psychophysics*, 19, 80–84.
Ross, J., & Di Lollo, V. (1971). Judgment and response in magnitude estimation. *Psychological Review*, 78, 515–527.
Schneider, B., Parker, S., & Upenieks, E. G. (1982). The perceptual basis of judgments of pitch differences and pitch ratios. *Canadian Journal of Psychology*, 36, 4–23.
Shepard, R. N. (1978). On the status of "direct" psychophysical measurement. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science* (Vol. 9, pp. 441–490). Minneapolis: University of Minnesota Press.
Stevens, J. C., & Rubin, L. L. (1970). Psychophysical scales of apparent heaviness and the size-weight illusion. *Perception & Psychophysics*, 8, 225–230.
Stevens, S. S. (1936). A scale for the measurement of a psychological magnitude: Loudness. *Psychological Review*, 43, 405–416.
Stevens, S. S. (1956). The direct estimation of sensory magnitudes—loudness. *American Journal of Psychology*, *69*, 1–25.
Svenson, O., & Åkesson, C. Å. (1966). *Fractional and multiple estimates in ratio scaling* (Report No. 202). Stockholm: University of Stockholm, Psychological Laboratories.
Taves, E. (1941). Two mechanisms for the perception of visual numerousness. *Archives of Psychology* (Columbia University), New York, No. 265.

EXPLORING MIND-MATTER INTERACTION: THE EFFECT OF MEDITATION ON THE BEHAVIOUR OF PHOTONS

Teodora Milojević and Mark A. Elliott
School of Psychology, National University of Ireland, Galway, Galway, Republic of Ireland

Abstract

Previous studies have reported an interaction occurring when meditation is employed in the well-known quantum physics, “double-slit” experiment (Radin et al., 2012). In this context, the act of focusing attention appears to collapse the photon wave-function causing them to behave in a more particle-like manner. This, in turn, decreases the intensity of the interference pattern – a series of alternating dark and bright bands produced by two or more waves interfering with each other. The present study employed a within-subjects design with 20 participants; 10 experienced meditators and 10 non-meditators. In an experimental “attention-toward” condition participants meditated on the double-slit, while in the control attention-away condition they did not attend, both in test sessions of 70, 20-second epochs. During both conditions, photon counts at the central maximum of the interference pattern were recorded and later compared in analysis. There was a significant difference in the intensity of the interference pattern at the central maximum recorded in the non-meditator and experienced meditator groups ($p = .005$). Experienced meditators, unlike non-meditators, induced a significant decrease in photon counts in the attention-toward condition relative to the non-attended condition ($p = .03$, $es = .19$). Finally, there was a strong positive correlation between meditation experience, expressed in the average number of hours spent meditating, and the observed decrease in photon counts ($r = .60$, $p = .04$). Contrary to the common belief that only matter has a unidirectional, first-order causal effect on cognition, the present results suggest that mental activities are capable of influencing physical systems. The findings and practical implications are discussed in light of previous research.

ABOUT THE BENEFITS OF CARRYING HANDLES ON MOVING BOXES

Friedrich Müller
Institute of Experimental Industrial Psychology (LüneLab),
Leuphana Universität, Lüneburg, Germany
<f.mueller@uni-lueneburg.de>

Abstract

Three experimental studies are introduced which were designed to explore the usefulness of scaling physical exertion using the Category Partitioning technique (CP) for the ergonomic optimization of lifting and carrying operations. Common to all experiments is that the participating subjects were each asked to pick up standardized moving boxes, to carry them, and place them on a table. Independent variables were the weight of the boxes and the ease of handling as some of the boxes had recessed grips, while others had to be manipulated without these aids. Dependent variables in all cases were CP-scaled physical exertion, but also heart rate (HR) in the second study. With very good reproducibility, the measurements show a linear relation between the weight of the boxes to be manipulated and the experienced exertion with a lower slope when the boxes are equipped with recessed grips. The scaling results correspond with heart rates registered in parallel in the 2nd experiment.

Even in highly technical, largely automated work environments, physically demanding work activities are often unavoidable. One example is the movement of baggage and general cargo when loading aircraft or conveyor belts at airports. In the process, each employee moves several tons of weight every day, often under difficult conditions in forced postures. In order to avoid overstraining and long-term consequences, employers, trade unions and the employers' liability insurance association are endeavoring to design work processes in such a way that health risks are avoided.

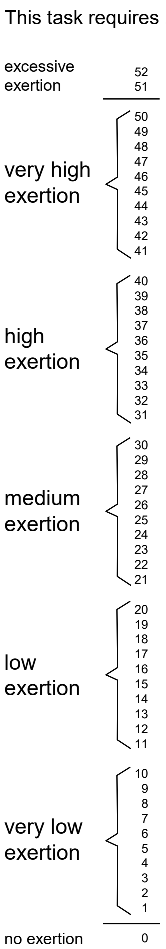
The basis for an ergonomic optimization of lifting and carrying operations, however, is the precise and reliable measurement of the experienced exertion. This requires a bias-free procedure that quantifies the absolute degree of exertion in each single case and moment, independent of the respective context and uninfluenced by comparisons with other activities and situations. The CP-Procedure, invented by Otto Heller (Heller, 1982) which is based on his orientation concept (Heller, 1990) seems to meet these requirements. CP scaled exertion was validated in extensive work trials on the bicycle ergometer by simultaneously recorded heart rates (Müller, Neely & Fichtl, 1995). It was shown, however, that the scaled data described the current strain, whereas the registered heart rates were influenced by previous work cycles. A more striking proof for the validity, however, was the observation that CP data obtained in single-stimulus experiments correspond exactly to scaling data from experiments with series of varying stress levels. And, after simple transformation, the CP measures agree well with data collected in otherwise identical experiments using a Category Ratio 20 Scale which was developed by Gunnar Borg (for a review see Borg, 1977). The results of long-term experiments in which we recorded the work and stress histories of young and older participants during seven hours of continuous work on a bicycle ergometer (Kakarot & Müller, 2014) confirm the earlier findings, that while both measurements are suitable to capture physical strain, heart rate is not as specific as CP-scaled exertion.

In the experiments presented below, we used the CP method to explore the effort experience of lifting and carrying commercial boxes. We were interested in the reproducibility of the scaling results, the relationship between weight and exertion, and the sensitivity of the scaling method to describe and quantify ergonomically effective design features such as, in this case, the inclusion of recessed grips in the side walls of the boxes.

Method

The Category-Partitioning Procedure (CP)

Fig 1. CP-Scale



The CP method consists of the name-giving scale (Figure 1) and a set of measurement rules as a prerequisite for an authentic quantitative description of perceived exertion. The CP-Scale contains five main category designations taken from everyday experience and formulated in common everyday language. Two additional categories ("no exertion" and "excessive exertion") anchor and allow to extend the scale, if a weight outside the personal experiences appears. Each of the colloquially named categories is divided by 10 ascending and equally spaced numbers, allowing fine differentiation within the given scale by a two-step modus first describing the magnitude by an appropriate verbal category, then finer grading by numerals within the category. This procedure increases the resolution of the scale and reduces so-called frequency effects as described by Parducci (1963).

Requirements for obtaining true-to-life measurements are an ecologically valid and isomorphic relation between the range of stimuli and the range of the scale used, on one hand and the subjects' previous experience of strenuous situations on the other hand. The memory stored experiences form the reference system for a metric perception and description of exertion (*mnestisch stabilisiertes Bezugssystem* in the sense of Witte, 1960). Incongruities of the scale extension and stimulus range invite orientation and adaptation processes that result in range effects.

Another requirement is to create an extraspective (in contrast to introspective) scaling mode. One should not refer to the personal feelings of the participants and prevent scrutinizing by not asking for "judgments", "guesses" or "impressions". Don't use phrases as "you cannot do wrong". But ask for facts like "how much exertion does this task require?" instead of "how exerted are you when lifting this box?"

Experiments and Results

1st Study: Lifting of boxes

To get a first impression about the feasibility, accuracy and replicability of the method we filled 8 standard moving boxes (675 x 365 x 340 mm) at even intervals with weights between 2.5 kg and 20 kg. We made sure that the weights (printer paper) were evenly distributed on the bottom surface of the boxes and could not slip. In permutated sequence, n = 19 participants (13 female and 6 male students) just lifted each of the boxes to a height of 72 cm and placed them on a table 50 cm away. Immediately after each lift, participants described the effort required by the task. The freely spoken instruction, tailored to each person, was: *"please place the boxes standing in the room in order on the table behind each one. Immediately after each lift, please describe how much exertion the task required. Use the scale provided as a reference for your description. Use the verbal expression first and then, for a more precise description, tell me the* *number within the category that best describes the required exertion. Then wait about 15 seconds and lift the next box*". After a one-hour break, the same persons lifted the boxes a second time in reverse order without knowing the sequence. The experimental procedure is summarized in Figure 2.

Fig. 2. Schematic representation of the experimental procedure. On the left side ground plan of the experimental room with 8 tables and 16 boxes. In the lower part the places of the experimenters.

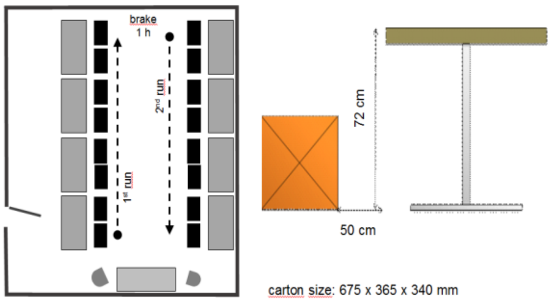
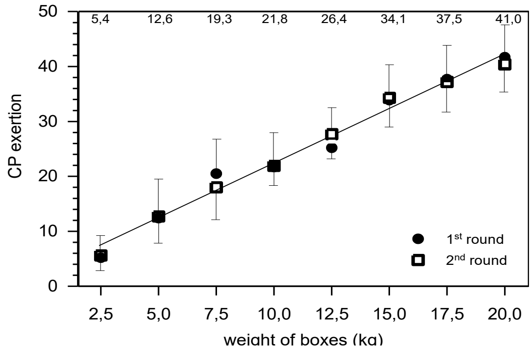


Fig. 3. Arithmetic means (M) calculated over n = 19 participants for the 1st (circles with SD pointing up) and 2nd round (squares with SD downwards). M for both runs as given at the top are basis for the regression line.



The scaling results presented in Figure 3 show a linear relationship between box weight and exertion required. The results of the two runs are almost identical ($r = 0{,}993$; $p < 0{,}000$). The arithmetic means calculated to the weights and the respective standard deviations differ only slightly. (The mean standard deviation in the 1st run is 5.3 scale units, in the 2nd run 5.4 scale units. The regression lines calculated for both runs (1st run: $y = 1.98 + 2.03x$; 2nd run: $y = 2.18 + 2.00\,x$) are almost identical. The regression line shown in Figure 3 ($y = 2.12 + 2.01x$) represents the scale values averaged over both runs, as indicated at the top of the graph.

2nd Study: Multiple lifting and carrying of moderately heavy boxes with and without handle

For each of the n = 19 students who had participated in the first experiments, plus one additional person, we estimated from the individual data the weight of the box whose lifting required about medium exertion (defined as CP = 25). For the N = 14 female participants the medium weight is M = 10.3 kg (SD = 0.62) and for the N = 4 male students M = 10.6 kg (SD = 0,40). All participants took part in the second study, conducted about 2 month after the first experiment. As sketched in Figure 4, moving boxes with the individually determined weights should then 10 times in succession and without interruption be lifted, carried 2.5 m to a table, briefly placed on the table and returned to the starting position. The entire process was timed and had to be completed within 190 seconds.

As independent variable the ergonomic handling option was introduced. A group of N = 10 participants manipulated the boxes using the standard handles attached at the sides of the boxes. For another group (n = 10), the handles were locked. CP-exertion was recorded after the first and the 10th sequence, respectively, together with the heart rate (HR) registered at these time points.

The results of this study are visualized in Figure 5. For actions without a handle, from the 1st to the 10th working sequence exertion increased from CP 18.6 to CP 30.9 and HR from HR 89.3 to HR 131.5. For the movements of the ergonomically designed boxes using the handles, CP-exertion and HR were significantly lower. Thus, exertion increased from CP 16.3 to CP 25.3 and HR from HR 91.9 to HR 118.3.

Fig. 4. Sketch of the experimental procedure in the 2nd study

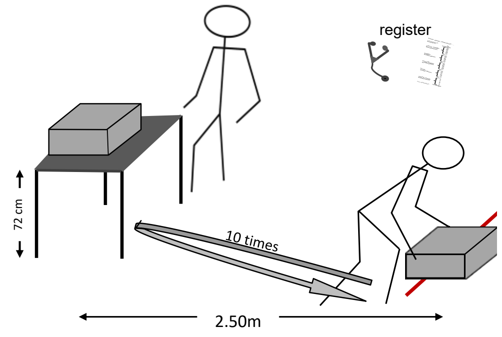
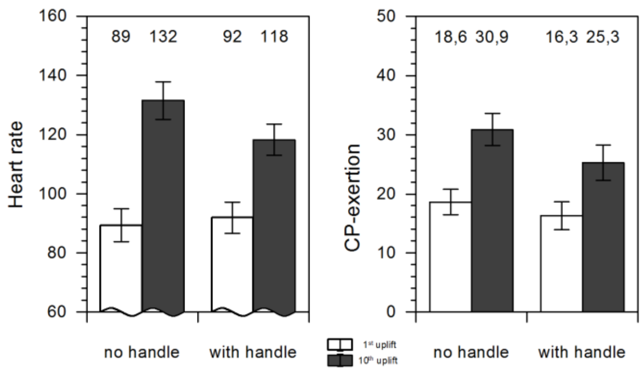


Fig. 5. HR and exertion at the beginning (white bars) and the end (black bars) of a 190 sec. lifting and carrying activity as a function of the ergonomic design of the manipulated boxes.



3rd Study: Lifting and carrying boxes of different weights with and without handle

In the third study, n = 30 university beginners (9 men and 21 women) were asked to carry in permuted sequence each of 20 prepared boxes (628 x 325 x 295 mm) with weights ranging from 1 to 20 kg over a distance of 2.5 m and to place them on a table. 10 of these boxes were carried using handles and 10 boxes had no handles. After each trial, the exertion was recorded. A stopwatch was used to verify that the runs lasted approximately the same amount of time between each other and additionally to ensure that the participants all had exactly three minutes of rest between the two runs. Before the trials began, participants were asked whether they were engaged in or practicing a physical activity and how often they do sport (from 1 = very rarely to 5 = very often).

Special attention was paid to the sequence of a total of 20 trials per participant. The requirements were that each participant lifts each of the 20 weight levels only once, of which 10 manipulations each done by using and without using a handle. The weights of boxes with and without handle should be distributed as evenly as possible between 1 kg (lightest) and 20 kg (heaviest weight), and the results of the permuted stimulus sequences should allow comparison in the sense of calculation a retest-reliability. These considerations led to the stimulus array shown in Table 1.

The results of this study agree quite well with the data obtained in the previous studies. The experienced exertion increased linearly with the weight of the cartons. As can be seen from Figure 6, the slope, however, is significantly flatter for the regression representing the more ergonomic handling (slope = 1.56) than for the lifts of the non-ergonomic boxes (slope = 1.72). For boxes of 20 kg, the relief of the more ergonomic compared to the boxes without support corresponds to an equivalent of about 2.5 kg. The repetitions of the experiment with a different permutation of the trials produced nearly identical results (Figure 7). A multiple regression model including the variables weight ($b = 1{,}639$, $\beta = 0{,}800$, $p < 0{,}000$), gender ($b = 4{,}302$, $\beta = 0{,}167$, $p < 0{,}000$), sport activity ($b = 1{,}822$, $\beta = 0{,}129$, $p < 0{,}000$) and handle ($b = 3{,}018$, $\beta = 0{,}128$, $p < 0{,}000$) calculates a corrected R^2 of 0,737.

Table 1. Lines 1 and 2 show the assignment of different weights to moving boxes with and without handles. Rows 3 and 4 list the boxes with the weight levels to be permuted for the lifts

in the first (1st stage) and second (2nd stage) halves of the test, respectively. The numbers correspond to the weights in kg.

no handle			3	4			7	8			11	12			15	16			19	20
using handle	1	2			5	6			9	10			13	14			17	18		

1st stage	1		3		5		7		9		11		13		15		17		19	
2nd stage		2		4		6		8		10		12		14		16		18		20

Fig. 6. Exertion when handling cartons with and without handle

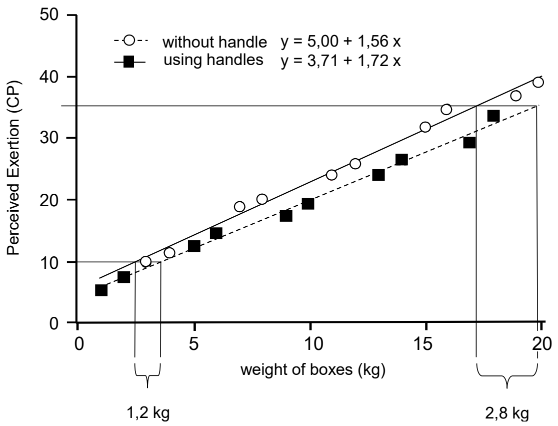
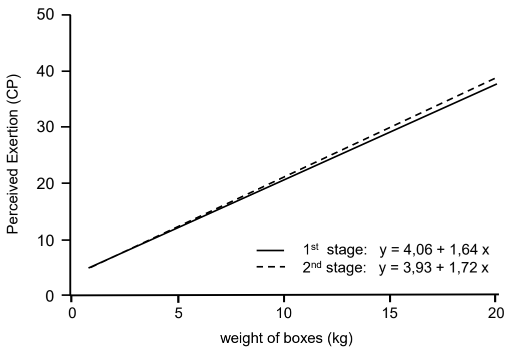
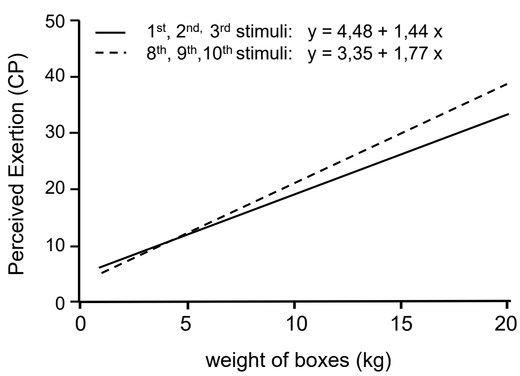


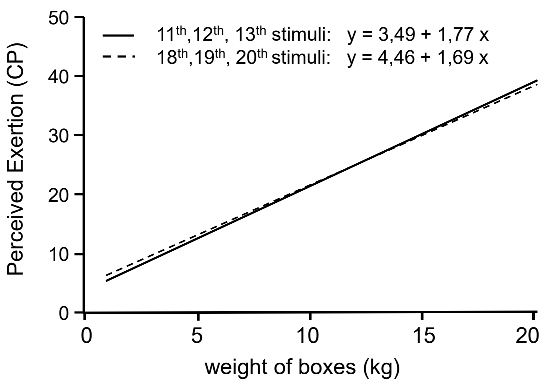
Fig. 7. Comparison of first and second stage



The regression lines shown in Figure 7 give the impression of a very good reproducibility and stability of the measurements of exertion. A closer inspection of the data, however, indicates a change in judgment behavior at the very beginning of the experiments, in such a way that the participants, in the sense of a range effect, further extend the judgments towards the upper range of the scale. This orientation or adaptation process is over after handling the third carton in each case (Figure 8). Since the scaling behavior stabilizes after the third lift, these orientation processes are hardly noticeable in the overall result. A possible cause for this in the sense of the CP technique undesirable effect could be that the participants, who were inexperienced in handling moving boxes, had not yet developed a reference system of box heaviness. In earlier experiments on scaling the pitch of instrument sounds (Müller & Steinbach, 1997), we observed that about three random samples from a population of related stimuli are sufficient to form an - initially however still unstable - reference system. We assume that no judging processes will occur in comparable experiments with professional movers, or if lay people had the opportunity to move a few boxes of different weights before the experiments begun.

Fig. 8. Regressions calculated for the first and last three lifts in the 1st stage (left picture) and the 2nd stage (right picture) respectively





Discussion

In summary, the CP technique has proven itself well in describing the exertion in lifting and carrying work. The effect of ergonomic measures on the experience of exertion is well and concordant described and quantified with the three studies. The results are stable between replicating trials, and the parallel recording of the heart rate gives indications of the validity of the scaling data. Scaling of exertion, ideally combined with posture analyses, can be an effective tool in optimizing workflows and improving working conditions in warehouses and transportation.

Acknowledgement

Dedicated to the memory of Gunnar Borg, who was a pioneer in the development of scaling methods for measuring exertion. With his work and dedication, he not only inspired research over decades, but also contributed significantly to the fact that psychophysical research has very successfully found its way into many areas of practice.

References

Borg, G. (Ed.) (1977) *Physical work and effort*. Wenner-Gren Center International Symposium Series (Vol. 28). England: Pergamon Press.

Heller, O. (1990) Scaling and orientation. In F. Müller (Ed.) *Fechner Day 90: Proceedings of the Sixth Annual Meeting of the International Society for Psychophysics* (pp. 52-57). University of Würzburg, Germany.

Heller, O. (1982) Theorie und Praxis des Verfahrens der Kategorienunterteilung (KU). In O. Heller (Ed.), *Forschungsbericht 1981*, Würzburg (Psychologisches Institut, Lehrstuhl für Allgemeine Psychologie), pp. 1-15.

Kakarot, N., Müller, F. (2014) Assessment of physical strain in younger and older subjects using heart rate and scalings of perceived exertion. *Ergonomics, 57*(7) 1052-67.

Müller, F., Steinbach, G. (1997) Range Effects in Pitch Scaling of Instrumental Sounds. In A. Preis & T. Hornowski (Eds.), *Fechner Day 97, Proceedings of the 13th Annual Meeting of the International Society for Psychophysics* (pp. 257-262). Poznan, Poland: International Society for Psychophysics,

Müller, F.; Neely, G & Fichtl, (1995). Scaling of Loudness, Perceived Exertion, and Pain Intensity: A Comparison Between the Category Partition (CP) and Category Ratio (CR-20) Scaling Procedures. In G. W. Neely (Ed.*), Category-ratio scaling of sensory magnitude in comparison with other methods*. Stockholm.

Parducci, A. (1963). The range-frequency-compromise in judgment. *Psychological Monographs, 77*(2), whole No 565.

Witte, W. (1960a). Experimentelle Untersuchungen von Bezugssystemen. *Psychologische Beiträge, 4*, 218-252.

SOME OBSERVATIONS ABOUT ECOLOGICAL MOMENTARY ASSESSMENT OF REAL WORLD EXPOSURE TO UNPLEASANT SOUNDS AND SMELL

Gregory Neely and Isabella Ström
Department of Psychology, Umeå University, SE-901 87 Umeå, Sweden

Abstract

Retrospective questionnaire methodology designed to survey individuals' exposures and experiences to real-world stimuli is susceptible various types of recall biases. Ecological Momentary Assessment (EMA) investigates individuals´ behaviour and experiences in real time by asking them to document their experience as soon as possible after exposure. This presentation reports some preliminary results from a study where a total of 114 participants in northern Sweden, comprised of 53 sound or smell hypersensitive persons (age: *M*=56, *SD*=12.7) and 61 non-hypersensitive individuals (age: *M*=61.1, *SD*=10.8), were surveyed through a questionnaire twice with a three-month period in between about exposure and experiences with unpleasant sounds and smells. During the three-month period, participates used a smartphone app to document incidents and experiences of unpleasant sounds between questionnaires. This presentation will discuss the similarities and differences observed between the results of the questionnaires and EMA as well as address methodological advantages and challenges presented by collecting EMA using a smartphone app system.

ORIENTATION BIAS IN AMBIGUOUS DOT LATTICES: CONTEXTUAL INFLUENCES AND ELECTROPHYSIOLOGICAL CORRELATES

Andrey R. Nikolaev[1, 2], Antonino Esposito[2, 3], Salvatore Gaetano Chiarella[2, 3], Antonino Raffone[3], Sergei Gepshtein[4], and Cees van Leeuwen[2, 5]
[1]*Department of Psychology, Lund University, Sweden*
[2]*Brain and Cognition Research Unit, KU Leuven, Belgium*
[3]*Department of Psychology, Sapienza University of Rome, Italy*
[4]*Salk Institute for Biological Studies, La Jolla, California, USA*
[5]*Center for Cognitive Science, TU Kaiserslautern, Germany*

Abstract

Perceptual bias reveals itself when generic representations of the world override effects of immediate stimulation. Perceptual bias may be elicited by prior contextual factors, such as visual working memory, and its signature brain activity could be detected by EEG. A perceptual bias for the vertical orientation from time to time overrides perceptual grouping by proximity in ambiguous dot lattices. We present an overview of our research on factors eliciting this orientation bias. We demonstrate how biased perception correlates with spontaneous EEG activity. We then show how temporal context elicits perceptual bias through top-down and bottom-up influences. We conclude that intrinsic preferences and the way they are elicited by the visual environment jointly shape our perception of the world. The ability to shift between the biased and unbiased modes provides flexibility to the visual system.

Visual perception does not start with light hitting the retina. Effects of stimulation depend on the broader visual environment and the state of the visual system at the moment the stimulus arrives. Acting in concert, these factors provide flexibility to visual perception. But the current understanding of the brain mechanisms behind these factors is far from complete. We review work of our group on how the intrinsic state of an observer and temporal contexts affect the perception of ambiguous visual stimuli (Esposito, Chiarella, Raffone, Nikolaev, & van Leeuwen, 2022; Nikolaev, Gepshtein, & van Leeuwen, 2016). We focus on the basic process of perceptual grouping by proximity. In multistable dot lattices, the stimulus components that are closest to each other tend to be perceptually grouped, in accord with the Gestalt law of proximity (Koffka, 1935). If, however, proximity allows for competing groupings, the perceptual outcome becomes a matter of probabilistic choice, which is conveniently measured in dot lattices (Kubovy, 1994). Dot lattices have allowed us to investigate the brain processes underlying the optimal and flexible processing of visual information.

Multistable dot lattices

Dot lattices are collections of regularly arranged dots that appear to be grouped into parallel strips (Kubovy, 1994). In Fig. 1A the four organizations are labeled **a**, **b**, **c**, and **d**. The shorter the distance between the dots in one of the directions, the more likely the dots group along that direction. According to the pure-distance law, the perceived organization of a dot lattice depends probabilistically on its aspect ratio, which is the ratio of the two shortest inter-dot distances: the ratio between **b** and **a**.

In all of the studies presented below we used rectangular dot lattices. The lattices were presented at four orientations, in which **a** was rotated counterclockwise from the horizontal for 22.5°, 67.5°, 112.5°, or 157.5° (Fig. 1A). Even though all orientations were oblique, the orientations 22.5° and 157.5° were close to horizontal, and the orientations 67.5° and 112.5° were close to vertical. The dot lattices were presented for 300 ms and were followed by a response screen. This screen consisted of four response icons, each containing a line parallel to one of the four likely organizations of the just-presented lattice. Participants reported the perceived orientation of dot lattices by clicking on one of the response icons. We asked participants to report the first orientation they perceived after the stimulus onset, emphasizing that the task had no correct answer.

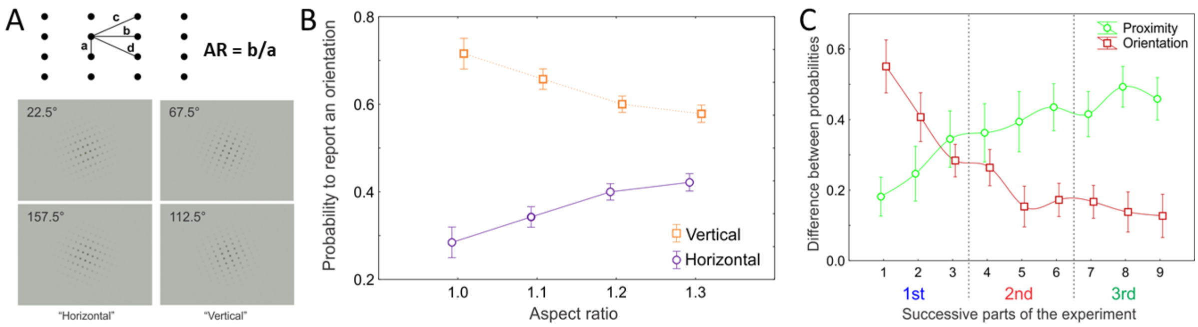


Fig. 1. **A**: (top) The perceived grouping depends on the aspect ratio (AR) of the dot lattice, which is the ratio of the two shortest inter-dot distances: along **a** (the shortest distance) and along **b** (the second shortest). In this case, since **b** is longer than **a**, vertical grouping is highly likely. (bottom) Four orientations of dot lattices: two of them are close to horizontal and two are close to vertical. **B**: Probabilities of the "vertical" and "horizontal" responses for each AR. The preference for vertical is largest for AR = 1.0, at which the lattices are most ambiguous, and decreases as the AR increases. **C**: Evolution of orientation and proximity responses during an one-hour experiment. The curve "Proximity" represents the difference between probabilities of responses consistent and inconsistent with the proximity principle. The curve "Orientation" represents the difference between vertical and horizontal response probabilities. "1st, 2nd, 3rd" at the lower abscissa indicate three parts of the experiment. The error bars are the standard errors across participants. Figures 1-3 are adapted from Nikolaev et al. (2016).

Vertical bias in perception of dot lattices

Dot lattices allow subtle manipulation of perceptual ambiguity. When the aspect ratio is close to one, stimulus support for any orientation is low and ambiguity is high. In this case, perception of orientation depends on perceptual bias. Perceptual bias is a product of the intrinsic state of the observer (Gepshtein & Kubovy, 2005). It can be understood as an intrinsic variable in the visual system that determines which percept is going to dominate in ambiguous situations (Chopin & Mamassian, 2011). Perceptual bias fluctuates slowly and stochastically, having dynamics similar to fluctuations of spontaneous brain activity. Therefore, Gepshtein and Kubovy (2005) proposed that the bias is driven by spontaneous activity. To test this hypothesis we looked for an association between spontaneous EEG activity and perceptual bias in dot lattices (Nikolaev et al., 2016).

In that study, we first analyzed the reported orientations of the dot lattices at four ARs: 1.0, 1.1, 1.2, and 1.3. We found that, in addition to the expected dependence of proximity groupings on AR, participants more frequently reported vertical than horizontal groupings. This is evidence for a bias for vertical orientation in perception of dot lattices (Claessens & Wagemans, 2008). The vertical bias is most prominent for the most ambiguous dot lattices (AR=1.0) (Fig. 1B). When stimulus support is weak, intrinsic factors guide perception, consistent with the notion that perception of ambiguous stimuli is the result of a tradeoff between intrinsic and extrinsic perceptual factors (Chopin & Mamassian, 2011; Gepshtein & Kubovy, 2005).

Remarkably, both the bias and proximity grouping changed systematically over time during the experiment: the effect of proximity on perception increased over time, while the effect of bias decreased (Fig. 1C). This pattern of results was the outcome of perceptual learning: the visual discriminability of small differences in inter-dot distances improved over time, whereas the dependence of responses on orientation bias declined. These effects changed much faster in the first third of the experiment than in later parts, suggesting specific processes in the initial, fast and saturating stage of perceptual learning, involving higher levels of visual perception (Sagi & Tanne, 1994). Although the bias decreased by the end of the experiment, it never disappeared completely, compensating for the stimulus ambiguity throughout the experiment.

To further explore the role of orientation bias in initial learning, we looked for serial dependencies in participants' responses. We investigated them using Lempel-Ziv complexity, a measure of randomness in binary time series. We computed Lempel-Ziv complexity for sequences of horizontal and vertical responses (orientation series), as well as for sequences of responses according to and against the proximity principle (proximity series), converting them to sequences of zeros and ones. We found that Lempel-Ziv complexity of the orientation series was lower than in the shuffled data, indicating serial dependency (Fig. 2A). The low complexity value indicates that the underlying process has a deterministic component. In the proximity series we found no serial dependency. To further investigate the serial dependency, we examined the repetitions of successive identical responses. We scored sequences of identical responses according to their length, e.g., three consecutive horizontal responses was scored as a sequence of length three. We found that the sequence lengths of actual data differed from the shuffled data throughout the experiment, though the difference was greater in the first part (Fig. 2B). The length of vertical sequences associated with biased responses was considerably longer than of horizontal sequences. Thus, the dynamic structure of orientation responses is characterized by temporal persistence.

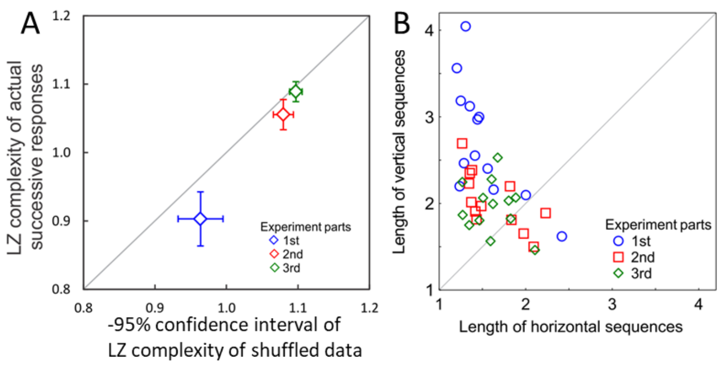


Fig. 2. **A**: Lempel-Ziv complexity for successive horizontal and vertical responses. The complexity of actual data is plotted against the lower bound of the confidence interval of mean Lempel-Ziv complexity in distributions of shuffled data for three parts of the experiment. The error bars are the standard errors across participants. **B**: Lengths of sequences of same responses for three parts of the experiment. Each data point represents one participant in one part of the experiment.

Prestimulus EEG alpha activity predicts perceptual bias

Since we sought to test the hypothesis that orientation bias is characterized by spontaneous brain activity, we recorded EEG from 256 electrodes. We focused on the EEG alpha activity, which dominates the spontaneous brain activity. Alpha activity reflects top-down processes of context and expectation, which may influence the neuronal excitability of sensory cortex (Jensen, Bonnefond, & VanRullen, 2012). Previous research has only investigated the relationship between prestimulus alpha activity and perceptual sensitivity in simple detection or discrimination tasks: low sensitivity was typically associated with high prestimulus alpha power (e.g., Wyart & Tallon-Baudry, 2009). The relationship between alpha activity and perceptual biases has not been studied previously.

We obtained alpha power over the parieto-occipital regions during one second *before* the presentation of dot lattices. We found that prestimulus alpha power was lower for vertical than for horizontal responses in the first third of the experiment, but not in later parts (Fig. 3A). Thus, prestimulus alpha activity predicts orientation bias when it is the dominant factor in response choice (Fig. 1C). No dependency on alpha activity was observed in proximity-based responses.

Next, we computed the Lempel-Ziv complexity of prestimulus alpha activity by converting its values in the trial sequences into binary series: each trial was coded as zero or one depending on whether its alpha power was below or above the mean power of that part of the experiment. The Lempel-Ziv complexity of the actual data was lower than that of the shuffled data only in the first part of the experiment (Fig. 3B). There was also a positive correlation between the complexities of behavioral responses and alpha power, suggesting that orientation preferences and alpha activity are driven by the same intrinsic dynamics. Finally, we related alpha activity to the sequence length of identical responses. Here we considered only sequences of lengths one, two, and three due to the insufficient number of longer sequences. We calculated the average alpha power across all trials of a sequence. We found a difference in alpha power between trials with vertical (biased) and horizontal (unbiased) responses only for sequences of two and three trials (i.e., for periods longer than 3 s), but not when responses were switched from one trial to another (Fig. 3C). Thus, prestimulus alpha power predicts biased responses only during long intervals. These findings support the hypothesis that lasting brain states accompany the elicitation of orientation bias (Gepshtein & Kubovy, 2005).

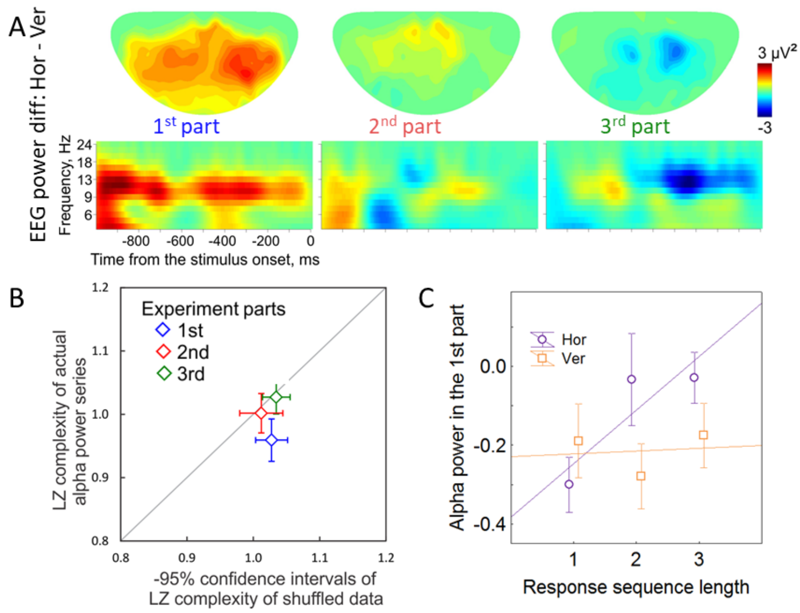


Fig. 3. Prestimulus EEG alpha power in three parts of the experiment. **A**: (top) The topography of alpha power differences between "horizontal" and "vertical" responses over the parieto-occipital brain areas. (bottom) Time-frequency plots of the power differences during 1 second before presentation of a dot lattice. **B**: Lempel-Ziv complexity of prestimulus alpha power in successive trials. **C**: Prestimulus alpha power for three sequence lengths of successive vertical and horizontal responses in the first part of experiment. The error bars are the standard errors across participants.

The effects of temporal context on orientation bias

Learning over the time course of the experiment (Fig. 1C) indicates that the perception of dot lattices depends on visual working memory. Therefore, lasting sequences of biased responses (Fig. 2) may be explained by preservation of representations of past dot lattices stimuli in working memory. Memory load and content can control in a top-down fashion whether proximity or bias dominates perception. Moreover, past stimuli can modulate perception in a bottom-up fashion through perceptual priming: for example, a dot lattice in vertical orientation may enhance detection of vertical dot lattices in the subsequent trials. To investigate the effects of top-down and bottom-up temporal context on orientation bias, we conducted two studies in which we manipulated the load and content of working memory during perception of the dot lattices, as well as the orientation of a prime stimulus presented shortly before the dot lattice (Esposito et al., 2022).

Top-down modulation of orientation bias

We designed a dual task experiment, in which we first presented participants with an array of four Gabor patches and asked them to memorize their orientations (Fig. 4A). Then, participants performed the dot lattice task, as in (Nikolaev et al., 2016). After that, one Gabor patch was presented in the same or different orientation as the initial configuration, and participants reported whether a change had occurred. To manipulate memory load, we presented all four patches in the memory array in the same orientation in the low-load condition, or two by two patches in the same orientation in the high-load condition. To manipulate the memory content, we considered a congruent condition when the memory array had the same orientation as the proximity-based one in the following dot lattice, and an incongruent condition when the memory array had a different orientation.

We found that both memory load and content modulated the orientation bias in perception of dot lattices. Participants were more likely to perceive dot lattices according to dot proximity in the low-load condition and according to bias in the high-load condition (Fig. 4B, C). We also found that the horizontal memory content decreased the vertical orientation bias, and this effect existed only in the low-load condition (Fig. 4D). Thus, the availability of free memory resources as well as the quality of memory representations reduce orientation bias, suggesting that in multistable perception, both memory capacity and content modulate the bias.

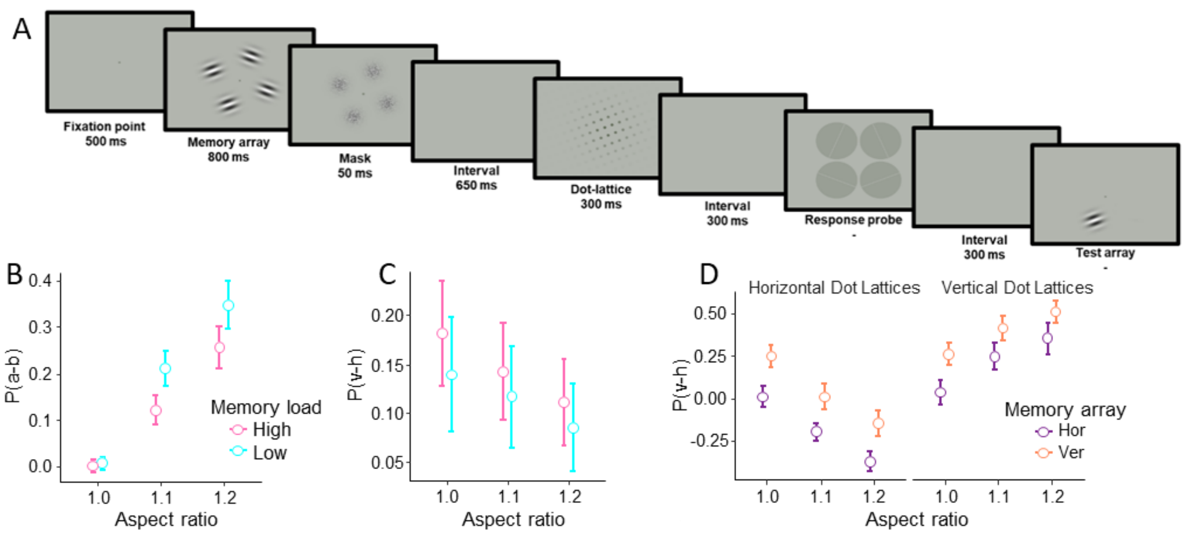


Fig. 4. The effect of memory load on orientation bias. **A**: Sequence of events of a single trial of the dual (memory and orientation judgment) task. **B**: Proximity-based index P(a-b) in memory load conditions. Proximity grouping interacts with memory conditions. **C**: Vertical-preference index P(v-h) in memory load conditions. Orientation bias is higher in the high-load condition, regardless of AR. **D**: Vertical-preference index P(v-h) after memorizing the vertical and horizontal arrays in the low-load condition. Orientation bias decreases when the horizontal array is stored in memory. The error bars are standard errors across participants. Adapted from Esposito et al. (2022).

Bottom-up modulation of orientation bias

We devised another task, in which participants determined the perceived orientation of a dot lattice immediately after exposure to a prime stimulus for 50 ms (Fig. 5A). The prime either had a prominent orientation (Gabor patch), which could be congruent or incongruent with the dot lattice, or the prime was a neutral noise patch. We found that, compared to the neutral prime, both horizontal and vertical primes increased vertical responses, i.e., biased perception of dot lattices (Fig. 5B, C). This effect was different from the top-down effects on the perception of dot lattices reported above. Oriented primes induce biased perception, but do not control it as oriented memory arrays do.

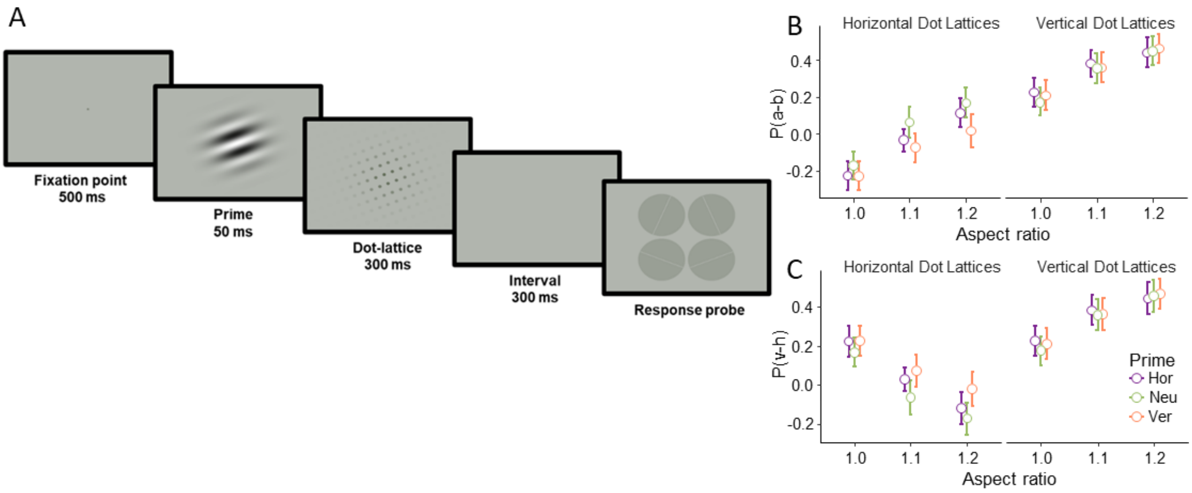


Fig. 5. The effect of priming on orientation bias. **A**: Sequence of events of a single trial of the priming task. **B**: Proximity-based index P(a-b) for the three primes. Proximity sensitivity is higher for neutral than for vertical primes for horizontal dot lattices. **C**: Vertical-preference index P(v-h) for the three primes. For horizontal dot lattices, orientation bias is higher after vertical than after neutral primes. The error bars are the standard errors across participants. Adapted from Esposito et al. (2022).

Conclusions

Perception of multistable dot lattices is guided by a tradeoff between the stimulus properties and the biases imposed by the observer. The tradeoff results in alternation of lasting episodes where the stimulus interpretation is based on either proximity or bias. This alternation correlates with spontaneous fluctuations of brain activity, indicating that perceptual biases reflect intrinsic states of the observer. However, past perceptual experience can modulate the intrinsic state. Bottom-up priming triggers the biased interpretations unselectively, leading to the rapid, biased resolution of visual ambiguity. By contrast, the top-down influences through working memory allow for more subtle and detailed representations of ambiguous stimuli. Together these factors support the flexibility of perception.

References

Chopin, A., & Mamassian, P. (2011). Usefulness influences visual appearance in motion transparency depth rivalry. *Journal of Vision, 11*(7), 18.

Claessens, P. M., & Wagemans, J. (2008). A Bayesian framework for cue integration in multistable grouping: Proximity, collinearity, and orientation priors in zigzag lattices. *Journal of Vision, 8*(7), 33 31-23.

Esposito, A., Chiarella, S. G., Raffone, A., Nikolaev, A. R., & van Leeuwen, C. (2022). Perceptual bias contextualized in visually ambiguous stimuli. *(subm.)*.

Gepshtein, S., & Kubovy, M. (2005). Stability and change in perception: spatial organization in temporal context. *Experimental Brain Research, 160*(4), 487-495.

Jensen, O., Bonnefond, M., & VanRullen, R. (2012). An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends in cognitive sciences, 16*(4), 200-206.

Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Hartcourt, Brace, and World.

Kubovy, M. (1994). The Perceptual Organization of Dot Lattices. *Psychonomic Bulletin & Review, 1*(2), 182-190.

Nikolaev, A. R., Gepshtein, S., & van Leeuwen, C. (2016). Intermittent regime of brain activity at the early, bias-guided stage of perceptual learning. *Journal of Vision, 16*(14), 11.

Sagi, D., & Tanne, D. (1994). Perceptual learning: learning to see. *Current Opinion in Neurobiology, 4*(2), 195-199.

Wyart, V., & Tallon-Baudry, C. (2009). How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. *Journal of Neuroscience, 29*(27), 8715-8725.

THE PRECISION OF FECHNERIAN INTEGRATION, QUANTIFIED

Lance Nizami and Claire S. Barnes
Independent Research Scholars, Bedford, MA 01730
<lancenizamiphd@comcast.net>

Abstract

A sensory stimulus evokes sensation; a just-noticeable difference (jnd) in that sensation delimits the so-called jnd in stimulus intensity. Hence, cumulating the intensity jnds upwards from the absolute stimulus-detection threshold (i.e., the just-perceptible stimulus) allows calculation of sensation growth. But counting jnds one-by-one has proven impractical; for more than a century, therefore, jnd-count was estimated algebraically through Fechnerian integration, which incorporates the Weber fraction. However, Fechnerian integration assumes infinitely-small jnds, an absurdity, one whose effect remains unexamined. Here, we pursue a neglected 60-year-old insight that models the jnd as a finite-difference equation, and we arrive at a converging infinite series of integrals that incorporate the Weber fraction, that yield the jnd count, and whose first term is the Fechnerian integral. We can now scrutinize the precision of the Fechnerian integral.

A sensory stimulus of physical intensity I evokes a sensation $F(I)$. The sensation could (for example) be loudness, saltiness, an odor, or a touch on the skin. F is usually assumed, based upon intuition, to be a monotonically increasing function of I. Another assumption is that we may ignore any changing aspect of sensation that is not caused primarily by intensity change, for example, a heard change in a sound-waveform's "pitch", primarily correlated with waveform frequency. Sensation-growth $F(I)$ has never been known for any sense, but the possibility of quantifying it continues to fascinate psychologists, physiologists, physicists, and philosophers. Customarily, $F(I)$ was presumed to be constructible step-by-step, as follows.

Intensity (of stimulus, or sensation) as a sum of just-noticeable intensity differences

Presumably, minimal perceptible sensation *differences* could be added together to produce the magnitude of the sensation evoked by a stimulus of a quantified intensity. Imagine a positive integer j that identifies a particular intensity I_j that evokes a sensation F_j. An empirical just-noticeable sensation difference $\Delta F \; at \; F_j$ corresponds to a just-noticeable intensity difference $\Delta I \; at \; I_j$:

$$\Delta I \; at \; I_j \; is \; called \; \Delta\left(I_j\right) \quad (1a) \quad corresponding \;\; to \;\; \Delta F \; at \; F_j, called \; \Delta\left(F_j\right) \quad (1b)$$

There must be a value of I_j below which, for a given research subject, a just-noticeable intensity difference cannot be obtained. A detection threshold $I_{th} = I_1$ can be declared for a stimulus, through statistical analysis of an observer's ability to detect the stimulus in controlled laboratory conditions. From $I_{th} = I_1$, where $\Delta(I_{th}) = \Delta(I_1)$, n successive increments $\Delta\left(I_j\right)$ cumulate to produce the intensity I_{n+1}, which is reached from I_n by the increment $\Delta(I_n)$. Hypothetically, the sensation increments $\Delta\left(F_j\right)$ likewise cumulate from $F(I_{th}) = F(I_1)$, the sensation at the stimulus-detection threshold I_{th}, to produce the sensation $F(I_{n+1})$, reached from $F(I_n)$ by the increment $\Delta F(I_n)$.

Let us now assume that there exist smooth, continuous relations $\alpha(I)$ such that

$$\Delta(I_j) = \alpha(I_j) \quad for\ j \geq 1, j \in \mathbb{I}^+ \tag{2}$$

Clearly $\Delta(I_j)/\alpha(I_j) = 1$. Let n be the cumulative number of just-noticeable intensity differences ΔI between $I_{th} = I_1$ and I_{n+1}. Hence,

$$n = \sum_{j=1}^{j=n} \frac{\Delta(I_j)}{\alpha(I_j)} \tag{3}$$

Usually, due to practical limitations of laboratory data-collecting, the $\Delta(I_j)$ are not known for a particular research subject or stimulus (and they can differ from subject to subject and, for example, from intensity to intensity of a given kind of stimulus). To compensate for this ignorance, we employ calculus. In this (the historical) approach, we imagine the just-noticeable intensity difference to shrink indefinitely:

$$\lim_{\Delta(I_j)\to dI} \sum_{I_j, j=1}^{I_j, j=n} \frac{\Delta(I_j)}{\alpha(I_j)} = \int_{I_1}^{I_{n+1}} \frac{dI}{\alpha(I)} \quad \cong n \tag{4}$$

The upper bound of integration is I_{n+1}, in order to accommodate the uppermost increment, $\Delta(I_n)$. Of course, as Nutting (1908, p. 292) reminds us, the just-noticeable differences in sensation and their associated just-noticeable differences in intensity "are finite quantities and by no means infinitesimal increments approaching zero as a limit". We therefore have only an *approximation*,

$$n \cong \int_{I_1}^{I_{n+1}} \frac{dI}{\alpha(I)} \tag{5}$$

Equation (5) represents classic "Fechnerian jnd-counting". Examples of this abound (see Troland, 1930; Miller, 1947; Majernik & Kalužný, 1975; McBride, 1983; Hellman & Hellman, 1990; see also Marks and Florentine, 2011, particularly p. 23). There is, however, a more precise method, hidden in the literature in a rarely-cited paper by Dallos and Carhart (1963), from a method applied by a mathematician, Haldane (1932), who applied the method not to hearing-science but to genetics. Let us precede the details of Haldane's (1932) calculus by first explaining that Dallos and Carhart (1963) abbreviated it for hearing-science.

Dallos and Carhart (1963): from difference limen to finite-difference equation

Using DL as their abbreviation for "difference limen", the just-noticeable intensity difference, Dallos and Carhart (1963) note that various sets of published DLs fit satisfyingly to

$$DL = a/\big(1 + (b \times SL)\big) \tag{6}$$

where SL is the sensation level (SL) in decibels, such that $SL = 0$ is the listener's stimulus-detection threshold, and $\{a, b\}$ are curve-fitted parameters, with a having units of decibels and b having units of inverse-decibels. Dallos and Carhart now propose that we consider a given SL to be a sum of DLs. Particularly, Dallos and Carhart quantify the sum of the first n DLs in an intensity range above some particular starting SL, not necessarily $SL = 0$. Call this sum the level S_n, in decibels with respect to the starting level. Then we can define the DL at S_n, which is the $(n+1)^{th}$ DL above the starting SL, as $S_{n+1} - S_n = \Delta S$. Hence, the very first DL is $S_1 - S_0$, where $S_0 \equiv 0$ (regardless of whether the starting threshold for the DLs is actually $SL = 0$). Altogether,

$$\Delta S = a/(1 + bS_n) \tag{7}$$

Dallos and Carhart recognized this as a *finite-difference equation*. Specifically, it is one of the kind solved by Haldane (1932). Dallos and Carhart (1963) therefore repeated a few of Haldane's key equations, prefatory to applying them to Eq. (7). But Dallos and Carhart omitted crucial details, enough that readers may have lacked confidence in applying the Dallos and Carhart method. (As of March 2022, Dallos and Carhart (1963) had received just six citations in the scholarly literature, according to the conservative GoogleScholar citation-counter.) The rest of this paper supplies the missing details, and goes far beyond.

Haldane (1932): the count of the finite differences in the finite-difference equation

J.B.S. Haldane was a polymath (Burke, 2020) who, amongst other things, applied mathematics to genetics (see the recent biography by Subramanian, 2020). Haldane's algebra (Haldane, 1932) starts with the equality $x_{n+1} - x_n = k\phi(x_n)$, where $n \geq 0$ is an integer which is a *count*. He then defines $x_{n+1} - x_n = \Delta x_n$. This approach has an implied use in psychophysics; substituting $x_n = I$ into $\Delta x_n = k\phi(x_n)$, where I is stimulus intensity, yields $\Delta I = k\phi(I)$, which resembles a Weber Fraction of the form $(\Delta I/I) = k\phi(I)/I$. Now, the count-of-differences $n \geq 0$ for some number x_n becomes the number of just-noticeable intensity differences that add to produce a stimulus intensity I. The latter notion was pursued by Dallos and Carhart (1963), but their method was peculiar to their own assumptions. A broader approach is presented here.

Let us commence with Haldane's (1932) proof, as follows. Haldane sets the count-of-differences $n \geq 0$ to

$$n(x_0, x_n) = \frac{1}{k}\int_{x_0}^{x_n} w(x,k)\,dx \tag{8}$$

where

$$w(x,k) = \sum_{r=1}^{\infty} \frac{k^{r-1}}{r!} f_r(x) \tag{9}$$

Haldane does not specify the nature of the functions $f_r(x)$. A hint comes from Haldane and Jayakar (1963, p. 302): "Haldane's series can be obtained simply by Abel's method", referring to Abel (1881). Note well that Eq. (8) for n is exact, not approximate. From Eq. (8), $1 = \frac{1}{k}\int_{x_n}^{x_{n+1}} w(x,k)\,dx$. Recall from above that $x_{n+1} = x_n + \Delta x_n = x_n + k\phi(x_n)$. Haldane now lets $x_n = x$ and $\phi(x) = y$, so that $x_{n+1} = x + ky$. Haldane also defines $F_r(x) = \int f_r(x)\,dx$. Making the appropriate substitutions into $k = \int_{x_n}^{x_{n+1}} w(x,k)\,dx$, and bearing in mind Eq. (9), yields

$$k = \int_x^{x+ky} w(x,k)\,dx = \sum_{r=1}^{\infty} \frac{k^{r-1}}{r!} \int_x^{x+ky} f_r(x)\,dx$$
$$= \sum_{r=1}^{\infty} \frac{k^{r-1}}{r!} \big(F_r(x+ky) - F_r(x)\big) \quad (10)$$

Haldane then expresses $F_r(x+ky)$ as a Taylor series, by treating ky as the variable and x as the fixed point around which the Taylor series is valid. Defining $F_r^s(x)$ as the s^{th} derivative of $F_r(x)$,

$$F_r(x+ky) = F_r(x) + \sum_{s=1}^{\infty} \frac{(ky)^s}{s!} F_r^s(x) \quad (11)$$

From Eqs. (10) and (11), then, and using $f_r^s(x)$ for the s^{th} derivative of $f_r(x)$,

$$0 = -k + \sum_{r=1}^{\infty} \frac{k^{r-1}}{r!} \sum_{s=1}^{\infty} \frac{(ky)^s}{s!} f_r^{s-1}(x) \quad (12a)$$

The math up to this point requires "that $\phi(x)$ should have no zeros or singularities in the region over which it holds, that is to say between x_0 and x_n" (Haldane, p. 235). As Haldane (1932) notes, convergence of these series to non-infinite entities depends upon k being "small". Haldane and Jayakar (1963) imply that $k \le 1$ for the most accurate outcomes of the series.

In Eq. (12a), k has the respective powers $\big((r-1)+s\big) = 1, \dots, \infty$ for $(r-1) = 0, 1, \dots, \infty$ and $s = 1, \dots, \infty$. Expanding Eq. (12a), while recognizing that $f_r^0(x) = f_r(x)$, yields

$$0 = -k + \left(\frac{ky}{1!} f_1(x) + \frac{(ky)^2}{2!} f_1^1(x) + \frac{(ky)^3}{3!} f_1^2(x) + \frac{(ky)^4}{4!} f_1^3(x) + \cdots\right)$$
$$+ \frac{k}{2!}\left(\frac{ky}{1!} f_2(x) + \frac{(ky)^2}{2!} f_2^1(x) + \frac{(ky)^3}{3!} f_2^2(x) + \cdots\right)$$
$$+ \frac{k^2}{3!}\left(\frac{ky}{1!} f_3(x) + \frac{(ky)^2}{2!} f_3^1(x) + \cdots\right) + \frac{k^3}{4!}\left(\frac{ky}{1!} f_4(x) + \cdots\right) \quad (12b)$$

We can now assemble the terms that respectively multiply k, k^2, k^3, and k^4. As Haldane notes, Eq. (12b) can be satisfied while coincidentally solving for $f_r(x), r = 1, \dots, \infty$, by equating each power of k to zero. For example, for the terms that are multiplied by k itself,

$$0 = -k + \frac{ky}{1!} f_1(x) \quad (13a)$$

$$\therefore f_1(x) = \frac{1}{y} \quad (13b)$$

For the terms multiplying k^2,

$$0 = \frac{k^2y^2}{2!} f_1^1(x) + \frac{k^2y}{2!\,1!} f_2(x) \tag{13c}$$

$$\therefore f_2(x) = \frac{1}{y}\frac{dy}{dx} \tag{13d}$$

Similarly, we obtain

$$f_3(x) = -\frac{1}{2}\left(\frac{1}{y}\left(\frac{dy}{dx}\right)^2 + \frac{d^2y}{dx^2}\right) \tag{13e}$$

$$f_4(x) = 2\frac{dy}{dx}\frac{d^2y}{dx^2} + \frac{1}{y}\left(\frac{dy}{dx}\right)^3 \tag{13f}$$

The first author confirmed these solutions.

Finally: the count of the just-noticeable intensity differences

Recall now the purpose of all this algebra, namely, to quantify n, the number of just-noticeable intensity differences between the stimulus-detection threshold and a given stimulus intensity. From Eqs. (8) and (9),

$$\begin{aligned} n(x_0, x_n) &= \frac{1}{k}\sum_{r=1}^{\infty} \frac{k^{r-1}}{r!} \int_{x_0}^{x_n} f_r(x)\,dx \\ &= \frac{1}{k}\int_{x_0}^{x_n} f_1(x)\,dx + \frac{1}{2!}\int_{x_0}^{x_n} f_2(x)\,dx + \frac{k}{3!}\int_{x_0}^{x_n} f_3(x)\,dx \\ &+ \frac{k^2}{4!}\int_{x_0}^{x_n} f_4(x)\,dx + \cdots \end{aligned} \tag{14}$$

The $f_r(x),\ r = 1, \dots, 4$ are now available in terms of $y = \phi(x)$ from Eqs. (13b, 13d, 13e, 13f). Altogether, the count-of-differences in terms of $y = \phi(x)$ is, to four significant terms,

$$\begin{aligned} n(x_0, x_n) \approx &\frac{1}{k}\int_{x_0}^{x_n} \frac{1}{\phi(x)}\,dx + \frac{1}{2}[ln\,\phi(x)]|_{x_0}^{x_n} \\ &-\frac{k}{12}\int_{x_0}^{x_n}\left(\frac{1}{\phi(x)}\left(\frac{d\phi(x)}{dx}\right)^2 + \frac{d^2\phi(x)}{dx^2}\right)dx \\ &+\frac{k^2}{24}\int_{x_0}^{x_n}\left(2\frac{d\phi(x)}{dx}\frac{d^2\phi(x)}{dx^2} + \frac{1}{\phi(x)}\left(\frac{d\phi(x)}{dx}\right)^3\right)dx \end{aligned} \tag{15}$$

The second term lacks k, i.e., there is always an early term that lacks k in the series $n(x_0, x_n)$.

Recall that Haldane (1932) started with the difference $x_{n+1} - x_n = k\phi(x_n)$, where $n \geq 0$ is an integer *count*, and x_n is the n^{th} value of some continuous, quantifiable thing x. Haldane then re-phrased $x_{n+1} - x_n = k\phi(x_n)$ as $\Delta x_n = k\phi(x_n)$. Let us convert from Haldane's x to the I used in sensory studies; then $x_n = I$, and $\Delta I = k\phi(I)$, and x_0 is the stimulus-detection threshold I_{th}. Replace k by K. Now, Haldane's count-of-differences $n(x_0, x_n)$ (Eq. 15) is

$$n(I_{th}, I) \approx \frac{1}{K}\int_{I_{th}}^{I} \frac{1}{\phi(I)}\, dI + \frac{1}{2}[ln\,\phi(I)]|_{I_{th}}^{I}$$
$$- \frac{K}{12}\int_{I_{th}}^{I} \left(\frac{1}{\phi(I)}\left(\frac{d\phi(I)}{dI}\right)^{2} + \frac{d^{2}\phi(I)}{dI^{2}}\right) dI$$
$$+ \frac{K^{2}}{24}\int_{I_{th}}^{I} \left(2\frac{d\phi(I)}{dI}\frac{d^{2}\phi(I)}{dI^{2}} + \frac{1}{\phi(I)}\left(\frac{d\phi(I)}{dI}\right)^{3}\right) dI \qquad (16)$$

In terms of the Weber Fraction $(\Delta I/I)$, we have $\phi(I) = I \cdot (\Delta I/I)/K$.

Conclusion: scrutinizing the precision of Fechnerian integration

The first term of Eq. (16) is the Fechnerian integral, Eq. (5). We now have a test of the precision of Fechnerian integration, which has been used to infer sensation growth for more than a century. A second paper in this series will evaluate that precision in the course of counting jnds for specific data-based Weber-fraction equations taken from the literature.

References

Abel, N.H. (1881). *Oeuvres complètes de Niels Henrik Abel. Tome second contenant les mémoires posthumes d'Abel.* Nouvelle edition publié aux frais de l'État Norvegien par MM. L. Sylow et S. Lie. Christiana. Impr. de Grøndahl & Søn. [Follows after Oeuvres complètes, 2 vols. (ed. B. Holmboe), 1839, Christiania.]

Burke, P. (2020). *The polymath: a cultural history from Leonardo da Vinci to Susan Sontag*. New Haven, CT, USA: Yale University Press.

Dallos, P.J. & Carhart, R. (1963). Cumulation of DL's for intensity change at low sensation levels. *Journal of the Acoustical Society of America*, *35*, 848-855.

Haldane, J.B.S. (1932). On the non-linear difference equation $\Delta x_n = k\phi(x_n)$. *Mathematical Proceedings of the Cambridge Philosophical Society*, *28*, 234-243.

Haldane, J.B.S. & Jayakar, S.D. (1963). The solution of some equations occurring in population genetics. *Journal of Genetics*, *58*, 291-317.

Hellman, W.S. & Hellman, R.P. (1990). Intensity discrimination as the driving force for loudness. Application to pure tones in quiet. *Journal of the Acoustical Society of America*, *87*, 1255-1265.

Majerník, V. & Kalužný, J. (1975). Intensity of the acoustical signals as an information carrier in the auditory transfer channel. *Acustica*, *32*, 174-180.

Marks, L. E. & Florentine, M. (2011). Measurement of loudness, Part I: methods, problems, and pitfalls. In M. Florentine, A. N. Popper, & R. R. Fay (Eds.), *Loudness*. New York, NY, USA: Springer, pp. 17-56.

McBride, R.L. (1983). Psychophysics: could Fechner's assumption be correct? *Australian Journal of Psychology*, *35*, 85-88.

Miller, G. A. (1947). Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. *Journal of the Acoustical Society of America*, *19*, 609-619.

Nutting, P.G. (1908). The luminous equivalent of radiation. III. Subjective light. 12. Photometric sensibility. *Bulletin of the Bureau of Standards*, *5*, 285-308.

Subramanian, S. (2020). *A dominant character: the radical science and restless politics of J.B.S. Haldane*. New York, NY, USA: W. W. Norton & Company.

Troland, L. (1930). *The principles of psychophysiology. II. Sensation*. New York, NY, USA: Van Nostrand.

STANDARD DEVIATIONS VERSUS MEANS OF PSYCHOMETRIC FUNCTIONS FOR DETECTION OF FORWARD-MASKED PROBE-TONES: UNEXPECTED NON-MONOTONICITY

Lance Nizami and Claire S. Barnes
Independent Research Scholars, Bedford, MA 01730, USA
<lancenizamiphd@comcast.net>

Abstract

Psychometric functions were obtained from two highly experienced subjects for detection of Gaussian-shaped 2-kHz probe-tones of 1.25 ms equivalent rectangular duration, each occurring 3 ms after a 200-ms 2-kHz fixed-intensity "forward-masker" of intensity of 10-90 dB SPL. Our use of some low forward-masker intensities is rare in the literature. The means and standard deviations of the psychometric function increased together for forward-maskers ≥ 50 dB SPL. But for 10-40 dB SPL forward-maskers, means barely differed from nonmasked probe-detection-thresholds, whereas standard deviations doubled. This was unexpected, and begs a re-assessment of forward masking, implying a "central" rather than "peripheral" origin.

An auditory stimulus's detection threshold and its discrimination threshold (its just-noticeable intensity difference, at any given intensity) are both elevated when the stimulus is preceded by a stimulus of similar frequency composition, and greater duration and/or intensity. This "forward masking" has been studied for decades. Remarkably, its origin remains unclear.

This paper presents the last remaining data from a series of experiments on detection and discrimination thresholds for Gaussian-shaped tones. Gaussian shaping of a stimulus is advantageous; it yields the smallest amount of spectral spread ("splash") for a given duration (Gabor, 1946), with the stimulus's relative spectral density showing only a single, relatively narrow lobe. A brief Gaussian-shaped stimulus can therefore be well-located in both time and frequency, making it a good probe for forward masking. The already-published experiments established probe-detection thresholds with (Nizami and Schneider, 1999) and without (Nizami, 2004) forward-masking, and established the probe's just-noticeable intensity differences with (Nizami et al., 2002) and without (Nizami et al., 2001; Nizami, 2006) forward masking. The data support various models of threshold, discriminability, and forward masking (Nizami and Schneider, 2000; Nizami, 2003; Nizami, 2013).

Method

The precision of obtained psychometric functions was a key element of some of the aforementioned studies. Presently, psychometric functions were constructed from percentages-correct, obtained as follows. The experiment itself employed Gaussian-shaped 2-kHz probe-tones of 1.25 ms equivalent rectangular duration. (For the duration conversion from Gaussian standard deviation to equivalent rectangular duration, see, for example, Nizami et al., 2001). Each such probe-tone appeared 3 ms after a 200-ms 2-kHz forward masker, its intensity chosen from 10-90 dB SPL. Listening trials consisted of 100 two-interval/two-alternative forced-choices per trial block, during which forward-masker and probe-tone intensities were fixed. Across-blocks, probe-tone intensity *x* (dB SPL) was varied, altogether providing a set of percentages-correct for probe-tone detection for a given forward-masker intensity. Figure 1 schematizes the stimuli and the listening task.

Percentages-correct for each fixed forward-masker intensity were fitted to a psychometric function $P(x)$, a sigmoid that is the integral of a Gaussian probability-density function having a mean value μ, taken to be the detection threshold, and a standard deviation σ, expressing psychometric-function breadth:

$$P(x) = \frac{1}{2} + \frac{1}{\sigma\sqrt{8\pi}} \int_{-\infty}^{x} exp\ \frac{-(x-\mu)^2}{2\sigma^2} dx, \quad (1)$$

$$where \quad -\infty \leq x \leq \infty,\ \ 0.5 \leq P(x) \leq 1$$

Also employed was a masker-free condition, yielding the psychometric function whose mean value is the *absolute* (i.e., nonmasked) probe-detection threshold.

The research subjects were two male college graduates, one in his twenties and one in his thirties, chosen because they were patient and thus had become highly accomplished in such listening tasks. Each subject sat wearing TDH-49 headphones in a double-walled soundproof chamber (ambient intensity ≈ 0 dB SPL). All stimuli were given only to the right ear, to avoid possible cross-ear effects (see Nizami, 2019).

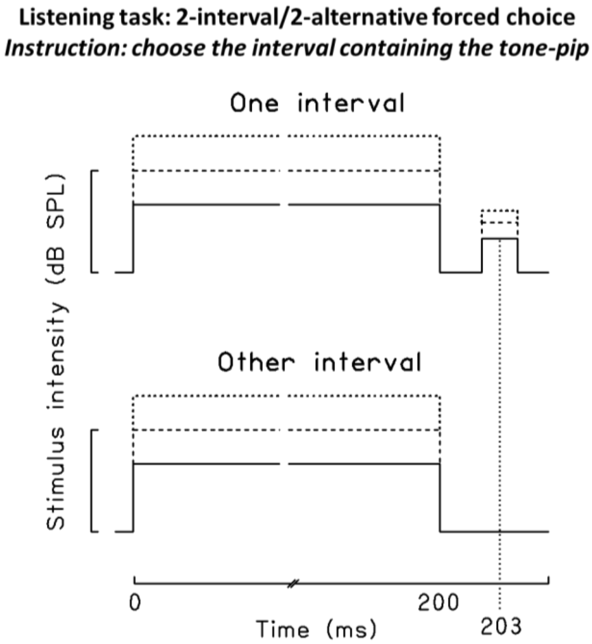


Fig. 1. Schematic of the stimuli and the listening task. The Gaussian-shaped probe-tone peaks at 203 ms in one of the two intervals (varied randomly). Dashed lines indicate different forward-masker intensities across trial blocks, and the consequent change in probe-tone intensity to maintain its detectability.

Predictions

The present study has special value because it uses some *low* forward-masker intensities, which are rare in the literature. Other investigators may have presumed that whatever did not affect the mean value could not affect the standard deviation.

Indeed, is there a model of how μ should relate to σ? Figure 2 schematizes the hypothetical effect of the cochlear mechanical nonlinearity upon the distributions of probe-tone intensities involved in probe-tone detection (Schairer et al., 2003):

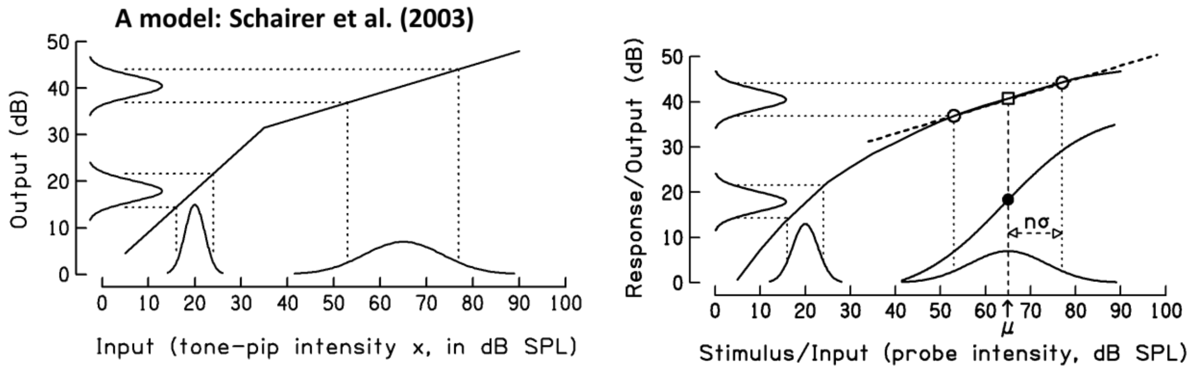


Fig. 2. (Left-hand side) The hypothetical relation of a presumed-constant Gaussian output distribution to a distribution of stimulus intensities, input to an idealized cochlear mechanical nonlinearity (whose units are decibels of peak displacement of the basilar membrane). (The units of the probability density functions are probability density, not decibels. Imagine, therefore, a probability density z-axis rising perpendicularly out of the page from (0,0), with the probability-density functions being projections upon the plane of the graph.) (Right-hand side) Similar to the left-hand-side, but with an *actual* cochlear mechanical nonlinearity (after Nuttall and Dolan, 1996, guinea pigs, and Rhode and Recio, 2000, chinchillas). Let us assume that the Gaussian-shaped input distribution is the Gaussian probability-density function that is integrated to make the psychometric function for probe-detection. The solid dot ● indicates the centroid of the psychometric function (corresponding to μ), and the open square □ is the respective locus on the cochlear nonlinearity. A convenient measure of the breadth of the psychometric function would be an even-numbered positive multiple of σ, say $2n\sigma$, where n is a positive integer. Its putative corresponding points on the cochlear nonlinearity are marked by the open circles ○. For simplicity, however, we will use σ itself as indicating psychometric-function breadth. (The units of the psychometric function are percentage correct, not decibels. Imagine, therefore, a percentage-correct w-axis rising perpendicularly out of the page from (0,0), with the psychometric functions being projections upon the plane of the graph.)

Probe-tone-detection threshold generally increases with the forward-masker's intensity (numerous sources). According to Schairer et al. (2003), then, and given the *actual* cochlear-mechanical nonlinearity (Fig. 2, right-hand-side), as μ increases with greater forward-masker intensity, σ will also increase, in a way that is nonlinear, smooth, and monotonic.

Results

Figure 3 shows the psychometric functions obtained.

Analysis

Figure 4 shows the standard deviation σ versus the associated mean value μ, for each psychometric function. The standard deviation σ of the psychometric function for the forward-masked 2-kHz Gaussian-shaped probe-tone does, indeed, monotonically increase with the mean value μ, *for forward-masker intensities > 50 dB SPL*. For lower forward-masker intensities, however, a more-complicated, non-monotonic relation of σ to μ occurs. To get the gist of the relation, we performed a trend analysis. Figure 5 illustrates the fit of a six-powers polynomial. We began with a polynomial of three terms altogether (the first three terms in Fig. 5), and then progressively added additional powers until there were six. Of all the polynomials, the six-powers polynomial offered the most-convincing fit, as judged by eye.

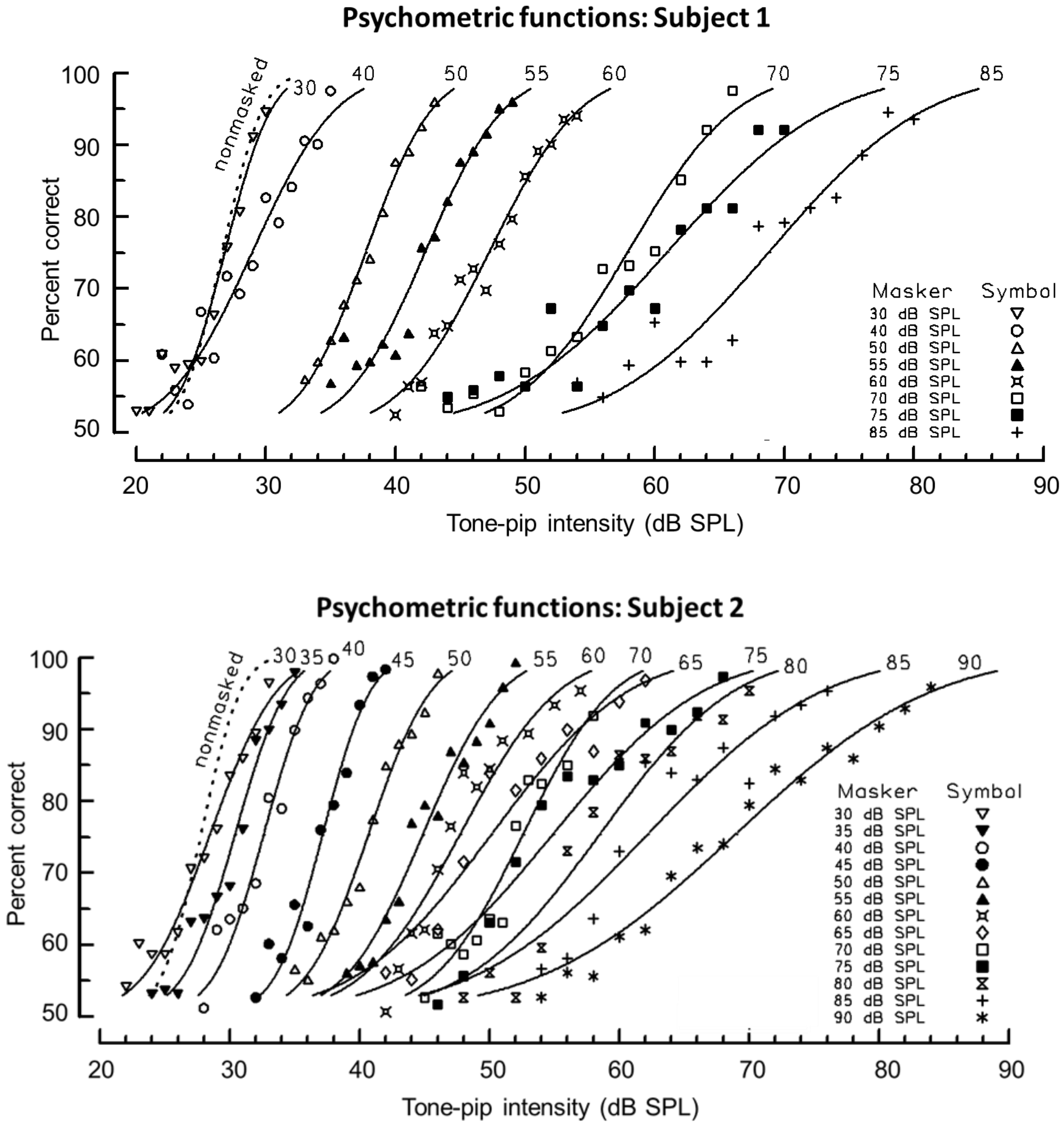


Fig. 3. Psychometric functions for the two subjects, for different forward-masker intensities.

Discussion and Conclusions

Our unexpected results for low-intensity forward maskers suggest some factor that is “central” (i.e., closer to the brain) rather than “peripheral” (i.e., the model of Schairer et al., 2003).

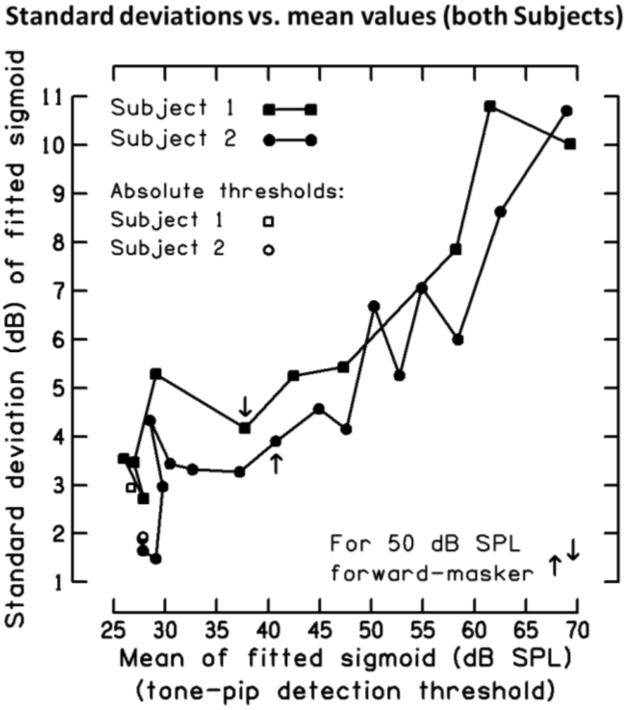


Fig. 4. Standard deviation σ versus mean value μ for each psychometric function of Fig. 3.

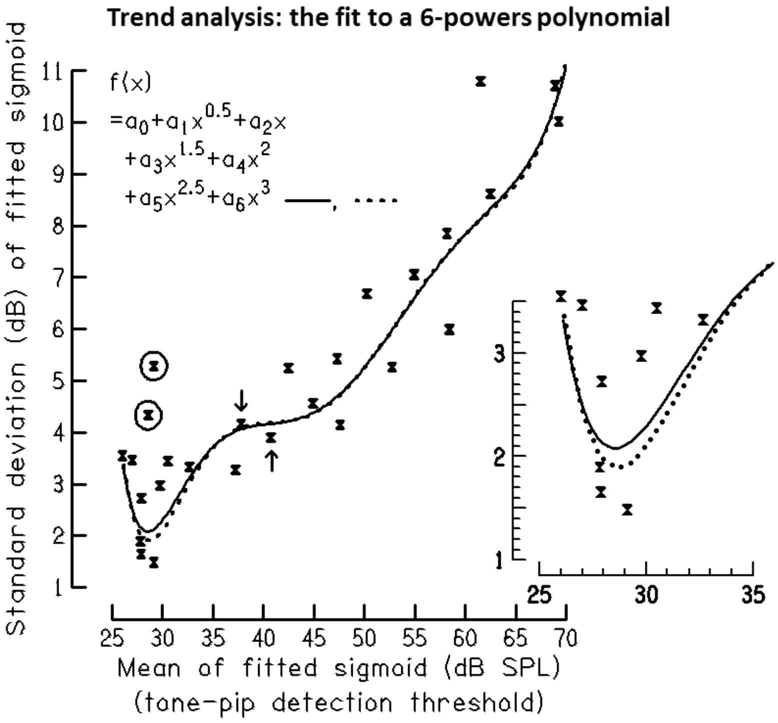


Fig. 5. The fit of a six-powers polynomial to the pooled data points of Fig. 4 (inverse-squares weighted). The arrows indicate the data for the 50 dB SPL forward masker. The two circled data-points seem unusually large, but momentarily excluding them (dashed lines) hardly changed the fit of the curve to the data.

Consider, for example, backward *detection* masking (unlike backward *recognition* masking; Bland & Perrott, 1978). In backward detection masking, the probe-tone *precedes* the masker; nonetheless, the probe-tone's detection threshold is elevated, but typically for a masker-probe time gap of no more than 20 msec. Although *forward* masking might partially reflect adaptation of the primary afferent auditory neurons (e.g., Harris & Dallos, 1979), backward masking may originate closer to the brain (e.g., Watanabe & Simada, 1971). All of this, and a large literature on "iconic memory" (e.g., Cowan, 1984), suggests a corruptible short-term central auditory memory, whose effective duration may resemble the "integration window" proselytized by Oxenham (2001) and others.

References

Bland, D.E. & Perrott D.R. (1978). Backward masking: Detection versus recognition. *Journal of the Acoustical Society of America*, *63*, 1215-1217.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin*, *96*, 341-370.

Gabor, D. (1946). Theory of communication. *Journal of the IEE*, *93*, 429-457.

Harris, D.M. & Dallos, P. (1979) Forward masking of auditory nerve fiber responses. *Journal of Neurophysiology*, *42*, 1083-1107.

Nizami, L. (2003). Afferent response parameters derived from postmasker probe-detection thresholds: 'the decay of sensation' revisited. *Hearing Research*, *175*, 14-35.

Nizami, L. (2004). Threshold vs. duration for Gaussian-shaped tone-pips of one to four periods duration. *Perceptual & Motor Skills*, *99*, 821-836.

Nizami, L. (2006). The intensity-difference limen for 6.5 kHz: an even more 'severe departure from Weber's law'. *Perception & Psychophysics*, *68*, 1107-1112.

Nizami, L. (2013). The human cochlear mechanical nonlinearity inferred via psychometric functions. *European Physical Journal – Nonlinear Biomedical Physics 1:3* (online).

Nizami, L. (2019). Simultaneous Dichotic Loudness Balance (SDLB): why loudness "fatigues" with two ears but not with one. *Attention*, *Perception*, *& Psychophysics*, *81*, 1624–1653.

Nizami, L., Reimer, J.F., & Jesteadt, W. (2001). The intensity-difference limen for Gaussian-enveloped stimuli as a function of level: Tones and broadband noise. *Journal of the Acoustical Society of America*, *110*, 2505-2515.

Nizami, L., Reimer, J.F., & Jesteadt, W. (2002). The mid-level hump at 2 kHz. *Journal of the Acoustical Society of America*, *112*, 642-653.

Nizami, L. & Schneider, B.A. (1999). The fine structure of the recovering auditory detection threshold. *Journal of the Acoustical Society of America*, *106*, 1187-1190.

Nizami, L. & Schneider, B.A. (2000). The periodicity of forward-masked auditory pip-detection thresholds, predicted from the output power of the auditory filter during ringing. *Hearing Research*, *144*, 168-174.

Nuttall, A.L. & Dolan, D.F. (1996). Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig. *Journal of the Acoustical Society of America*, *99*, 1556-1565.

Oxenham, A.J. (2001). Forward masking: Adaptation or integration? *Journal of the Acoustical Society of America*, *109*, 732-741.

Rhode, W.S. & Recio, A. (2000). Study of mechanical motions in the basal region of the chinchilla cochlea. *Journal of the Acoustical Society of America*, *107*, 3317-3332.

Schairer, K.S., Nizami, L. Reimer, J.F., & Jesteadt, W. (2003). Effects of peripheral nonlinearity on psychometric functions for forward-masked tones, *Journal of the Acoustical Society of America*, *113*, 1560-1573.

Watanabe, T. & Simada, Z. (1971). Auditory temporal masking: an electrophysiological study of single neurons in the cat's cochlear nucleus and inferior colliculus. *Japanese Journal of Physiology*, *21*, 537-549.

DISCREDITING THE "DISCREDITING" OF PSYCHOPHYSICS: H.K. BEECHER VERSUS THE HARDY-WOLFF-GOODELL DOLORIMETER

Lance Nizami and Claire S. Barnes
Independent Research Scholars, Bedford, MA 01730
<lancenizamiphd@comcast.net>

Abstract

In 1947, Hardy, Wolff, and Goodell achieved a psychophysics milestone: they built a putative sensation-growth scale, for skin pain, from pain-difference limens. Limens were found using the "dolorimeter", a device first made by Hardy & co. to evoke pain for pain-threshold measurements. Scant years later, though, H.K. Beecher (MD) discredited the pain scale – according to Paterson (2019), citing the historian Tousignant. Yet Hardy & co. receive approval in the literature. Intrigued, we scrutinized their methods, then Beecher's critiques, and Tousignant's history of threshold dolorimetry. Beecher decried dolorimetry as irrelevant, favoring clinical trials of pain relief. *But he failed to discredit dolorimetry.*

Writing in the journal *Body & Society*, Paterson (2019) emphasizes that pain is distinct from touch. Paterson also emphasizes the philosophy that *internal* pain, not to be confused with skin pain, is part of a homeostasis mechanism. Paterson (p. 118) notes that a scale for *skin* pain was constructed by Hardy, Wolff, and Goodell in the 1940's (citations below); but subsequently, "The Hardy–Wolff–Goodell scale was discredited in 1957 by the distinguished clinical anaesthesiologist Henry K. Beecher, who became Chair of the Committee on Drug Addiction and Narcotics and who established a Laboratory of Anaesthesia (Tousignant, 2014: 122)". No further details emerge. The cited Tousignant (2014) is a book chapter, derived from a lengthier peer-reviewed paper (Tousignant, 2011) which was a history of the use of a skin-pain-evoking device built and used by Hardy et al. – the very same device used to construct their "discredited" pain scale. But Tousignant denies "discrediting" that pain scale (personal communication, 1 April 2022). Nonetheless, Tousignant's (2011) review does indeed (and prominently) feature the above-mentioned Henry K. Beecher, MD. His numerous publications sometimes concerned military medicine, but not psychophysics. How, then, did Beecher discredit the psychophysics of Hardy et al.? Here, we describe their psychophysics, and its clinical use, and Beecher's attempts to dismiss it in favor of clinical *trials*. We begin with Hardy et al.'s pain thresholds.

Measuring the threshold for pain, with and without pain reduction by drugs

Establishing a pain-detection threshold

Hardy et al. (1940) wished to evoke controlled pain in human skin. Historically, the stimuli were "mechanical, chemical, electrical, and thermal" (Hardy et al., p. 649). Hardy et al. favored thermal, by radiation, thanks to (1) relatively easy construction of equipment, (2) accurate measurement of stimulus intensity, (3) "sharply defined" sensory thresholds (Hardy et al., p. 649), (4) manipulability of exposure duration and of skin condition, (5) ability to choose skin patches (regardless of texture) of different size, and (6) "The stimulus can be repeated in rapid succession without injury to the skin surface tested" (p. 649). Also (p. 649), "The sensation produced is sharp, a "bright pain", and is to be distinguished from an ache or deep pain".

Hardy et al. (p. 650) described the apparatus and method as follows: "The light from a 1000 watt lamp was focussed [*sic*] by a condensing lens through a fixed aperture onto the blackened forehead of the subject. The surface of the forehead to be tested was thoroughly blackened with India ink. This measure was taken to insure total absorption of the radiation, regardless of pigmentation of the skin, and to eliminate possible effects due to the penetration of the rays below the skin surface. The stimulus could thus be considered as purely thermal". Further (Hardy et al., p. 650), "The intensity of the radiation was controlled by means of a rheostat. Immediately in front of the lamp was mounted an automatic shutter, which was arranged to allow the radiation to pass through to the subject for exactly 3 seconds". Pain threshold was then established by exposing 3.5 cm^2 of blackened forehead-skin. If the subject felt no pain, then the light intensity was increased, the exposure being repeated after 30-60 sec. As Hardy et al. (p 650) note, "This procedure was followed until the subject just felt pain at the end of the exposure. This threshold pain was easily recognizable even by untrained subjects. The sensation was that of heat finally "swelling" to a distinct, sharp stab of pain at the end". A radiometer was placed where the subject's forehead had been, to measure the pain-threshold radiation intensity (expressed in gram calories per second per square centimeter, abbreviated gm. cal./sec./$cm.^2$). Remarkably, half of the thresholds agreed within 2% over re-tests (Hardy et al., p. 651); here and elsewhere, Hardy, Wolff, and Goodell were their own research subjects.

Measuring skin-pain-threshold elevation by analgesics

Hardy, Wolff, and Goodell immediately applied their pain-threshold-detection method to assessing "the action of analgesic agents". Their motivation: "No adequate method for assaying their effects on the pain threshold in man has been available. However, since the prime purpose of an analgesic drug concerns its action in man, it is desirable to measure accurately its effect on man's pain threshold" (Wolff et al., 1940, p. 659). Note well that Hardy, Wolff, and Goodell were doing classic *psychophysics* on healthy volunteers (themselves, using their above methods); they were not doing *pharmacology* with clinical patients. They established the "control" pain threshold; then, "an analgesic agent was administered [intramuscularly] and observations of the pain threshold were made at 10-minute intervals until the threshold had returned to the control level, that is, until all pain threshold-raising action had ceased. The height of the pain threshold-raising effect was expressed in per cent elevation above the control level" (Wolff et al., 1940, p. 659). The latter elevation was plotted versus time for a "time-action curve". These were obtained for opioids, with special attention to morphine sulphate; its time-action curve rose, peaked, and fell, reaching greater height (i.e., pain relief) with greater dosage. Supplementary doses prolonged the relief. Psychophysics had proven useful to the pain clinic.

Measuring skin-pain-threshold elevation by analgesics during "internal" pain

Hardy and colleagues then embarked upon a far more dangerous program. That is, "prolonged pain was introduced as a variable since in this way the action of morphine could be appraised more nearly in terms of its common therapeutic use" (Wolff et al., 1940, p. 672). Three methods were used: inflating a blood-pressure cuff to 200 mm Hg pressure on the upper arm, or "swallowing a catheter to which was attached a balloon which was distended with water when it reached the duodenum" (Wolff et al., p. 672), or clamping the trapezius and biceps muscles. All three treatments lasted 40 *minutes*, and each produced a different kind of "deep, aching pain". The actual administration of morphine sulphate and the pain-threshold measurements (radiation on the forehead, as above) proceeded thus: "After the control readings, which preceded the morphine injection, the painful procedure was begun: (1) 46 minutes before injection; (2) 1 minute after the injection; (3) 50 minutes after the injection; (4) 120 minutes after the injection. Pain-threshold readings were made every 10 minutes throughout the subsequent 6 to 7 hours" (Wolff et al., p. 672). Compared to the case *without* deep aching pain, whose time-action plot lasted 7 hours, the general finding was that 0.015 grams of morphine sulphate caused the time-action plot to remain the same for case (1), to peak sooner and lower but to decline faster for case (2), to generally shrink for case (3), and to peak at the same level but decline faster for case (4). (These changes are difficult to picture, but there is no space for pictures.) Hardy and colleagues had shown that analgesia could be quantified psychophysically.

Building a scale for pain from just-noticeable pain differences

The just-noticeable difference (jnd) in sensation (*difference limen*) is a standard psychophysical measure. In principle, jnds of warmth, heat, and pain are empirically obtainable. But Paterson (2019, p. 118) disparages jnds of pain as "an arbitrary pain intensity unit". Units imply a *scale*.

Herget et al. (1941) used the Hardy-Wolff-Goodell apparatus to obtain jnds. They employed a two-alternative two-interval method-of-limits, alternating 2 sec. light-exposures between the right and left halves of the blackened area of the forehead, keeping one side at a fixed intensity (0-35,000 $\times 10^{-5}$ gm. cal./sec./cm.^{2}) while increasing the illumination of the other side step-wise until the subject could just detect a difference between left and right. The research subjects were one or more of Herget et al. themselves (otherwise unclear). The jnds (ΔI) follow three stages, each consisting of a rise that decelerated to a plateau. Herget et al. (p. 651) labeled the stages for lower, middle, and high intensities respectively as "warmth" ("mild, pleasant, diffuse"), "heat" ("sharper and sometimes stinging"), and "pain" ("sharp, biting, and granular").

Using the same equipment, but a somewhat different method, Hardy et al. (1947) carefully obtained jnds for pain. Notably, "In the series of experiments with stimuli greater than 500 millical./sec./cm.^{2} [units of 0.001 gram cal./sec./cm.^{2}], considerable tissue damage was produced. For this reason, a second test area, the blackened volar surface of the forearm, was chosen. This area had the same pain threshold as the forehead and was more easily cared for when blistered" (Hardy et al., p. 1153). The discovered jnds (ΔI) were approximately constant for 220-320 millical./sec./cm.^{2}, then increased roughly monotonically over 320-680 millical./sec./cm.^{2}, the last illumination providing pain saturation. Hardy et al. found 21 jnds between threshold and saturation. They then adopted Fechner's classic postulate, that each jnd represents an equal sensation change. Thus, plotting jnd count from 0 to 21 versus the radiation intensity produced a pain-growth plot. It became a pain-growth *scale*, a psychophysics ideal, when Hardy et al. declared that any two adjacent pain jnds constituted one pain unit, the *dol*.

Hardy, Wolff, and Goodell encounter Henry K. Beecher

The Hardy-Wolff-Goodell apparatus was dubbed the "dolorimeter". Tousignant (2011) details the history of its clinical use; she reports that "By 1950, over twenty research teams had published data generated by a dolorimeter; it had entered laboratories across the United States, and traveled to Britain and Canada" (Tousignant., p. 147). This popularity was highest in the late 1940's during a "burst of commercial interest in synthetic analgesics" (Tousignant, p. 164). However, "From 1950, the number of articles reporting the evaluation of analgesic drugs with the dolorimeter – its most popular usage – leveled off. In 1953, the validity of dolorimetric analgesic tests was under attack in the pages of *Science*" (Tousignant, p. 148).

Indeed it was. The attacker was Henry K. Beecher MD. His offensive started in 1952, in a six-page salvo promoting clinical-trials research. Beecher (1952, p. 159) stated that "During work on pain in 1947, we were led to postulate that there is a fundamental difference in what can be learned in studying "natural" pain which arises in a pathological focus (disease or trauma are defined here as "natural" cause) from that produced experimentally (heat to forehead, pin pricks, electric shocks, or heat to teeth, pain deliberately produced with a tourniquet, and so on). The basis for this postulate had its beginning in our attempts to use the Hardy-Wolff-

Goodell technique". Such attempts had evoked some complaints. As Tousignant (2011, p. 153) notes, "Although Hardy, Wolff, and Goodell initially claimed that their instrument worked with untrained subjects, they later admitted the importance of practice and instruction (they used the term familiarization rather than training) to obtain more consistent pain thresholds". Further (Tousignant, p. 165), "Even at the height of its popularity, few denied that using the Hardy–Wolff–Goodell method required adjustments and compromises. It was found to lack sensitivity to the effects of weaker analgesics such as aspirin. A more significant issue was the difficulty in replicating the consistency in threshold values obtained by Hardy, Wolff, and Goodell".

Beecher (1952, p. 159) opined that "It requires little imagination to suppose that the sickbed of the patient in pain, with its ominous threat against his happiness, his security, his very life, provides an entirely different milieu (*and reaction*) than the laboratory, with its dispassionate and unemotional atmosphere" (original italics). Of course, we might dispute whether laboratory-induced pain involves an unemotional atmosphere. But back to Beecher (p. 159): "In the experimental pain experience, the relatively short duration of the stimulation and the experimental situation make the experience primarily one of pain sensation. We do not believe that the pathological pain situation with all the diffuse associations of illness, disease, and pain can be satisfactorily reproduced in the laboratory". Hardy and his co-authors had already evoked pain in themselves that was pseudo-pathological, lacking "the diffuse associations" (above). Beecher (1952, p. 160) further declared that "Since pain is almost always a consequence of disease or pathological trauma, the study of pathologic pain seems to us the more direct and logical approach to an understanding of the pain experience and its relief".

Hardy, Wolff, and Goodell respond to Beecher

Hardy, Wolff, and Goodell (1953, p. 165) responded promptly: "Pain *sensation* is far more difficult to investigate when an individual is extremely frightened, inattentive, obtunded, prostrated, "sick," or exhausted. On the other hand, these would be ideal circumstances for the assay of an agent designed to make the patient "more comfortable." The bedside method is the only one that will ultimately establish whether a given analgesic has a place in clinical medicine. On the other hand, the separately studied effects of an agent on the pain threshold, pain intensity, and reactions to noxious stimulation, local and general, are of vital interest to the investigator and therapist" (original italics). In other words, psychophysics could co-exist with clinical medicine. Beecher (1953a, p. 166) replied: "It is plain that the reaction of the man in a sickbed, where his pain may be a warning of disaster, will not be the same as the reaction of a well and comfortable man in the laboratory subject to a momentary pricking sensation", emphasizing that "it seems to us reasonable to separate pain on the basis of its origins and significance to the subject; that is, experimental or pathological". But the two pains could be endured simultaneously (Wolff et al., 1940), on which Beecher (1952, 1953a) had no comment.

Beecher's approach: precision through population

How Beecher measured pain

Later in 1953, Beecher re-stated his preferred approach to measuring pain: "Measurement of pain depends upon how much analgesic is required *to relieve the pain*. To be sure this is indirect, but no more so than determination of the acidity of a solution by the quantity of standard alkali used to neutralize it. We depend upon average pain, defined as the *average response elicited from 25 or more individuals in pain* (postoperative)" (Beecher, 1953b, p. 323; italics added). Tousignant (2011, p. 152) explains: "Beecher's subjects – postoperative patients – were asked to estimate the relief of their pain as "none, slight, moderate or complete"; later they would be asked only to choose between more than 50 percent relieved or not". Note well that psychophysical measurement of pain *threshold* of laboratory staff is now replaced by expressions of pain *relief* amongst a cohort of patients – a cohort, not only because of the crudeness of the data, but also because, amongst other things, "Sound design of the experiment requires that willing, cooperative, undistracted subjects be used in sufficient numbers to cancel out normal mood swings above and below par. The body's diurnal temperature swings, with their demonstrable effects on performance, also require controls" (Beecher, 1952, p. 160).

Tousignant synopsizes (2011, p. 172) the data analysis: "Statistical analysis, standardized observation procedures, and access to large amounts of subjects were essential given Beecher's conceptualization of the experimental subject as a collective one, whose main virtue was abundance rather than stability, certainty, or detachment. The precision and accuracy of the analgesic clinical trial did not rely on individual judgments of pain relief but on their aggregation". Of course, "more" is not "more credible"; a study's credibility depends upon the credibility of each outcome, not their sheer number. But Beecher (1952, pp. 160-161) declared that "Mathematical validation of any supposed [treatment-caused] differences is essential". This is ignorance. Math cannot unequivocally validate differences (e.g., Rozeboom, 1960; Dunnette, 1966; Abbott, 2013; Smith, 2018; Nizami, 2019; Kuorikoski, 2021); results deemed "significant" may fail attempts at replication, particularly in psychology (e.g., Yong, 2018).

Complications in clinical trials

Beecher (1952, p. 160) rationalized clinical trials, so: "We are concerned incidentally, of course, with simplicity. A method that can function with no apparatus other than a notebook and pencil is manifestly more desirable and more broadly useful, other things being equal, than one that requires complex and delicate apparatus which needs calibration by a well-trained physicist". But the replacement of psychophysical measures by clinical ones brought new complications. Now, the drug-giver had to be ignorant of what was given – whether it was the active ingredient or merely a placebo, used as a control. Placebos, too, caused complications: "We are interested in studying the pharmacology of a new drug. We try it out on a group of patients; a third to a half of this group will be relieved of their symptoms by a placebo; they react favorably to the syringe regardless of what it contains. Thus they dilute the significant data derived from the other half or two thirds of the group that react only to the drug contained in the syringe. We are not, in studying a new drug, interested in the pharmacology of syringes; we are nonetheless obliged to take into account the placebo reactors; *we must screen them out* …" (Beecher, 1952, p. 161; italics added). And stimulus presentation had to be randomized. Further, treatment outcomes could depend upon who administered the substances, and who recorded their effects!

Not surprisingly, then, "the analgesic clinical trial was more expensive, time-consuming, and difficult to coordinate than the dolorimetric method. Using simple interrogation – "fairly primitive questions" as one of Beecher's colleagues would later say – to obtain comparisons of pain-relieving efficacy required a large team of observers, subjects, and consultants to collect and manipulate information under appropriate conditions, and on a sufficiently large scale. Observers and statistical experts needed to be paid salaries and consultant fees; while recruiting subjects required the authority needed to access a clinical setting, coordinate clinical staff, and oversee patient treatment" (Tousignant, 2011, pp. 156-157). All this is true today.

Conclusions

Paterson (2019, p. 118) states that "The Hardy–Wolff–Goodell [pain] scale was discredited in 1957 by the distinguished clinical anaesthesiologist Henry K. Beecher". Paterson implicates Hardy et al.'s measurement of just-noticeable pain differences (pain jnds) in Beecher's critique.

But Dr. Beecher (MD) criticized Hardy et al. rather for their method of assessing pain mitigation by drugs, namely, the elevation of the pain-detection threshold, measured psychophysically using Hardy et al.'s *dolorimeter* – also used by Hardy et al., a few years later, for pain jnds. Beecher disparaged psychophysics in favor of clinical trials. But dolorimetry and clinical trials had already been recognized as different means to similar ends; Beecher merely made the dolorimeter a "straw man" and then burnt it down. Tousignant (2011, p. 166) offers a final perspective: "The dolorimeter fit into a gap between animal studies (convenient, relatively cheap and easy to standardize, but with limited translatability to humans) and clinical ones (perhaps more authentic, but much more difficult, and expensive, to standardize)".

References

Abbott, D. (2013). The reasonable ineffectiveness of mathematics. *Proceedings of the IEEE*, 101, 2147-2153.

Beecher, H.K. (1952). Experimental pharmacology and measurement of the subjective response. *Science*, 116, 157-162.

Beecher, H.K. (1953a). *Response*: Pain – controlled and uncontrolled. *Science*, 117, 166-167.

Beecher, H.K. (1953b). A method for quantifying the intensity of pain. *Science*, 118, 322-324.

Dunnette, M.D. (1966). Fads, fashions, and folderol in psychology. *American Psychologist*, 21, 343–352.

Hardy, J.D., Wolff, H.G., & Goodell, H. (1940). Studies on pain: a new method for measuring pain threshold: observations on spatial summation of pain. *Journal of Clinical Investigation*, 19, 649-657.

Hardy, J.D., Wolff, H.G., & Goodell, H. (1947). Studies on pain: discrimination of differences in intensity of a pain stimulus as a basis of a scale of pain intensity. *Journal of Clinical Investigation*, 26, 1152-1158.

Hardy, J.D., Wolff, H.G., & Goodell, H. (1953). Pain – controlled and uncontrolled. *Science*, 117, 164-165.

Herget, C.M., Granath, L.P., & Hardy, J.D. (1941). Thermal sensation and discrimination in relation to intensity of stimulus. *American Journal of Physiology*, 134, 645-655.

Kuorikoski, J. (2021). There are no mathematical explanations. *Philosophy of Science*, 88, 189-212.

Nizami, L. (2019). Information Theory is abused in neuroscience. *Cybernetics & Human Knowing*, 26, 47-97.

Paterson, M. (2019). On pain as a distinct sensation: mapping intensities, affects, and difference in 'interior states'. *Body & Society*, 25, 100-135.

Rozeboom, W.W. (1960). The fallacy of the null-hypothesis significance test. *Psychological Bulletin*, 57, 416-428.

Smith, G. (2018). *The AI Delusion*. New York, NY: Oxford University Press.

Tousignant, N. (2011). The rise and fall of the dolorimeter: pain, analgesics, and the management of subjectivity in mid-twentieth-century United States. *Journal of the History of Medicine and Allied Sciences*, 66, 145-179.

Tousignant, N. (2014). A quantity of suffering: measuring pain as emotion in the mid-twentieth-century USA. In: R. Boddice (Ed.), *Pain and Emotion in Modern History*. New York, NY: Palgrave Macmillan, pp. 111-129.

Wolff, H.G., Hardy, J.D., & Goodell, H. (1940). Studies on pain: measurement of the effect of morphine, codeine, and other opiates on the pain threshold and an analysis of their relation to the pain experience. *Journal of Clinical Investigation*, 19, 659-680.

Yong, E. (19 Nov 2018). Psychology's replication crisis is running out of excuses. *The Atlantic*.

TEGHTSOONIAN'S LAW IS NOT

Lance Nizami and Claire S. Barnes
Independent Research Scholars, Bedford, MA 01730
<lancenizamiphd@comcast.net>

Abstract

In a 1962 lecture, S.S. Stevens summarized exponents of sensation-growth power functions discovered by him and his associates. In 1967, E.C. Poulton tabulated some of those exponents, along with the respective ratios of the highest stimulus intensity to the lowest stimulus intensity. In 1971, Professor Robert Teghtsoonian graphed the Stevens exponents versus the logarithms of the stimulus-intensity ratios, finding a hyperbolic plot. It transpired to fit well to a hyperbolic curve, now called "Teghtsoonian's Law", which Teghtsoonian derived from Stevens' power function by relating Stevens' exponent to the stimulus-intensity ratio and the corresponding sensation-range ratio, the latter being assumed constant. Here, we emphasize that stimulus-intensity units are arbitrary, making comparisons of stimulus-intensity ratios meaningless. We also compile, from literature not associated with Stevens, sets of Stevens exponents and their respective stimulus-intensity ratios. Teghtsoonian's Law does not fit.

S.S. Stevens (1962) laid out exponents of power functions for empirical plots of sensation versus stimulus intensity, found by himself and associates. From that, psychophysicist E.C. Poulton tabulated "the [Stevens power-function] exponents and stimulus ranges of the 21 out of 27 sensory dimensions for which adequate experimental data have been published" (Poulton, 1967, p. 312). Poulton found that for those 21 data points, "the tau coefficient of rank correlation between size of exponent and geometric stimulus range is -.60 (p< .001). This means that 36 percent of the variance of the tabulated exponents is accounted for simply by the geometric range of stimulus variables used in determining them" (Poulton, p. 312).

Poulton's finding was noted by Professor Robert Teghtsoonian, a founder of the International Society for Psychophysics. Teghtsoonian (1971, p. 72) wrote: "It follows from the simple form of the power law

$$\psi = \alpha\phi^n \tag{1}$$

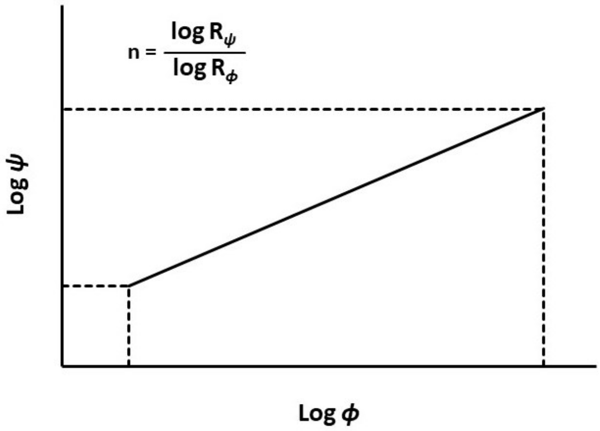
that

$$n = \frac{\log R_\psi}{\log R_\phi} \tag{2}$$

where R_ϕ is the ratio of the greatest to the smallest stimulus intensity, and R_ψ is the ratio of corresponding sensory magnitudes". Figure 1 (below) encapsulates these concepts. Teghtsoonian (p. 72) continued: "If the range of log judgments provided by *S*s is nearly constant, we may substitute *K* for $\log R_\psi$. Then

$$n = \frac{K}{\log R_\phi} \text{ ."} \tag{3}$$

Teghtsoonian (1971, p. 72) declared that "An estimate of this constant K may be obtained by determining the value of $n \log R_{\phi}$ for each of the 21 experiments [Poulton, 1967], and calculating the mean". That mean was 1.53. Teghtsoonian substituted 1.53 into Eq. (3), then plotted a curve using $\log R_{\phi}$ as abscissa and n as ordinate, along with Poulton's actual data points. Figure 2 reproduces the work. A correlation is visible, and was re-emphasized in subsequent papers (e.g., Teghtsoonian, 1973, 2012, 2018; Teghtsoonian & Teghtsoonian, 1978, 1997; Teghtsoonian et al., 1981). It became known as "Teghtsoonian's Law". It is widely cited (GoogleScholar) and is still being praised in the literature (e.g., Laming, 2009; Ward, 2017).



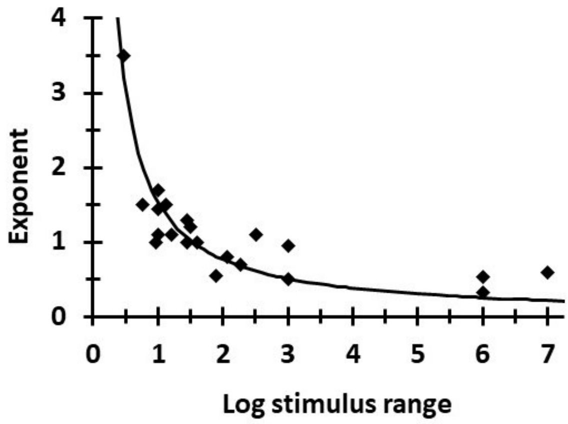


Fig. 1. Stevens power function, with associated quantities. Fig. 2. Teghtsoonian's Law (curve).

What Teghtsoonian concluded

Teghtsoonian (1971, p. 72) proposed that "the ratio of the greatest to the smallest possible sensory magnitude is approximately constant for all perceptual continua, and that variation in power law exponents among continua reflects variation in the ratio of the greatest to the smallest stimulus intensity to which S is responsive (a ratio which is sometimes called the dynamic range)". Further, Teghtsoonian (1973, p. 5) emphasized that "Since an investigator designing a scaling experiment will generally seek to present as wide a range of stimulus intensities as is possible commensurate with his subjects' comfort, the result may be to select a range of values that is nearly the same constant proportion of the dynamic range (on a log scale) for all continua. The resulting values of $\log R_{\psi}$ will be the same constant proportion of its maximum and hence will tend to be constant". The value of $\log R_{\psi}$ was later amended from 1.53 to 1.64 (Teghtsoonian et al., 1981), based upon maximum physical effort using a bicycle ergometer, but it was reinstated as 1.53 in later work (e.g., Teghtsoonian & Teghtsoonian, 1997).

The Teghtsoonians bolster Teghtsoonian's Law

Teghtsoonian and Teghtsoonian (1997) had subjects quantify "force of handgrip, heaviness, loudness, electric shock, and sniff vigor" (Teghtsoonian & Teghtsoonian, p. 723). For each subject and for each kind of stimulus, the experiment first established "a range [of intensity] that was as wide as possible while being "acceptable and reasonably comfortable"" (Teghtsoonian & Teghtsoonian, p. 724). These were the equivalents of R_{ϕ} in Eq. (3). The equivalents of n were then found by quantifying sensation growth by fitting it to Stevens power function. Across-subject medians $(n, \log R_{\psi})$ provided five data points, one for each kind of stimulus. Altogether, they conformed closely to Teghtsoonian's Law.

Scrutinizing Teghtsoonian's Law

Why a Ratio of Intensities for Defining Dynamic Range?

The Teghtsoonians never dwelt on why a *ratio* of intensities was chosen to define a dynamic range. A ratio is not the only possibility. Suppose that R_ϕ is a ratio of intensities, a *geometric* range. A logarithm of a ratio of intensities, $\log R_\phi$ in Eq. (3), is a difference of logarithms of intensities, an *arithmetic* range for intensities expressed using a logarithmic intensity scale. But one might use a non-logarithmic, Cartesian intensity scale to quantify differences, such that the dynamic range would be defined as the difference between the highest and lowest intensities, i.e., the arithmetic Cartesian range. Here is a conundrum. As Poulton (1967, p. 312) explained: "The stimulus range is defined geometrically [by some] as the ratio of the largest variable to the smallest variable. This is clearly unsatisfactory from a theoretical point of view, since halving the smallest variable makes little difference to the arithmetic range; yet it has as great an effect on the geometric range as doubling the largest variable, which practically doubles the arithmetic range". True, so far. Poulton (p. 312) continued: "But it is not possible to define the stimulus range arithmetically, since it then depends upon the sizes of the units used in each physical dimension, which are arbitrary". True again. Indeed, the units of stimulus intensity represent "an arbitrary historical decision" (Weiss, 1981, p. 432).

Poulton (1967) did not delve further. He implied that the geometric range, Teghtsoonian's ratio R_ϕ, does *not* depend upon units. This may represent the oft-taught idea that units "disappear" from ratios of quantities having the same units, i.e., that units "divide out", making ratios unitless. But herein lies a trap. Poulton may have thought that any ratio whose units "divide out" could be compared to any other ratio whose units divide out, *even if the respective ratios started with different units*. That is, Poulton seemingly assumed that a ratio of two weights, for example, could be compared to a ratio of two luminances. The same assumption is implicit in Teghtsoonian's work. The assumption is wrong. *Algebraic operations, such as taking a logarithm, or raising to a power, do not eliminate the nature of the things that are operated upon, and neither does division of one such thing by another.*

The Persistence of Units in Ratios: Implications for the Weber Fraction

The above concepts have wide scope. Consider the following. Observe the ratio at the heart of psychophysics, the Weber ratio or "Weber fraction", $\Delta I/I$, composed of the just-noticeable-difference ΔI at intensity I, divided by I. Myers (1982, p. 210) noted that "it has been customary to consider that the minimum value of the [Weber] ratio is an index of the maximum resolving power, or sensitivity to change, of a given sense mode, and that the resolving power of different sense modes may be compared by comparing the numerical values of their Weber ratios. Textbooks often present rankings of different sense modes based on their Weber ratios". The Weber ratio is a ratio of intensities having the same units – units chosen by convention. *By convention*, because *there are no unique units that characterize intensity*.

Fechner (1877) famously related intensity to sensation, labeling intensity as β. Regarding β, Boring (1928, pp. 444-445) demanded "what in the world is β, the 'intensity' of a stimulus? Is β energy, or force, or pressure, or surface tension, or power, or some other physical quantity, and is it the quantity measured in some aspect of an external stimulus-object, or in some other aspect, or at the sense-organ, or in the sense-organ, or at the receptor, or somewhere else? In any case where Weber's Law [$\Delta\beta/\beta = constant$] might apply, there are several possibilities for the measurement of β, and often the different possibilities in the same case are for quantities that could not be related linearly".

What are "quantities that could not be related linearly"? Myers (1982, p. 205) explained: "In brightness scaling, luminance is the commonly used measure of stimulus intensity. Regardless of the units used, luminance is a linear function of luminous power density per unit solid angle per unit projected area of the target surface in the direction of the viewing eye. That is, given a constant spectral energy distribution, it is directly proportional to stimulus energy. Nevertheless, rather than designating the intensity of visible electromagnetic radiation by its power density, one could instead specify the intensity of its electric field. This is common practice in radio-frequency field measurements, and one does occasionally see this method used in the measurement of light. The point is that luminance is a linear function of power (rate of energy flow) but is proportional to the *square* of electric field intensity (when acting in a constant field impedance, which we are assuming here). (One could also scale the stimulus in units of magnetic field intensity, with similar consequences.) It is mainly a matter of custom that luminance is referenced to power rather than to electric or magnetic field intensity" (original italics and brackets). Regarding units in acoustics, root-mean-square (RMS) *acoustic power* (sometimes called intensity) is energy per unit time, and is proportional to RMS *pressure* squared; *acoustic energy* (per m^2 of area) thus equals intensity multiplied by the stimulus duration (Hartmann, 1998, p. 31).

The Weber Fraction Under Units that are Nonlinearly Related: an Example from Myers

Consider what differences in units of intensity imply for the magnitude of the Weber fraction. The Weber fraction is a ratio of quantities that have identical units; the magnitudes of the quantities in the ratio depend upon those units, and hence, so does the magnitude of the Weber fraction. An example for psychoacoustics came from Myers (1982, pp. 210-211): "Take the classic results of Riesz (1928) on auditory intensity discrimination for a 1000 Hz tone at 80 db above threshold. Eighty db is an acoustic field power of approximately .01 microwatts per square centimeter (μW/cm^2). Riesz found a Weber ratio $\Delta I/I = .05$, or in units of acoustic field power, $\Delta I/I = (.0005\ \mu\text{W}/\text{cm}^2)\ /(.01\ \mu\text{W}/\text{cm}^2) = .05$. Nevertheless, if the stimulus had been rescaled in acoustic field pressure in Newtons per square meter (N/m^2), the corresponding values would be $\Delta I/I = (.0051\ \text{N/m}^2)\ /(.204\ \text{N/m}^2) = .025$. Transforming from units of acoustic power to acoustic pressure halves the Weber ratio. Or suppose we did the calculation in decibels, a logarithmic transformation of acoustic power: $\Delta I/I = .21\ \text{dB}\ /80\ \text{dB} = .0026$".

Intensity Ratios Under Changes of Units: Summary of Implications

Comparing ratios is not meaningful when two ratios involve different units. Likewise, neither is comparison of any mathematical transformation of those ratios, such as their logarithms. Yet Poulton (1967, Table 1) listed stimulus-intensity ratios (here denoted R_ϕ) for stimuli within and across sensory modalities, as if comparison was justified, and Teghtsoonian continued the error, plotting Stevens n versus $\log R_\phi$ (Teghtsoonian, 1971; Teghtsoonian, 1973, Fig. 1; Teghtsoonian & Teghtsoonian, 1997, Fig. 1; Teghtsoonian, 2012, Figs. 1 & 2).

Why does a meaningless correlation seem so convincing?

Given that Teghtsoonian's Law accounted for "over 87% of the variance in the reported exponents" (Teghtsoonian, 1971, p. 73), how could it be inappropriate? Regarding the data of Fig. 2, Teghtsoonian (1971, p. 73) states that "The points for binaural loudness, monaural loudness, vocal effort (the autophonic response), and vibration intensity are plotted using values of $\log R_\phi$ which are just half those shown by Poulton. In expressing a given number of decibels as a ratio, Poulton has not taken account of the fact that in these four cases the reported exponents were calculated for ranges of stimulus pressure or amplitude; in such cases the number of decibels [composing the stimulus range] is 20 $\log R_\phi$, not 10 $\log R_\phi$ as Poulton assumed". Supposedly, Poulton had assumed units of RMS intensity (i.e., *power*) for the decibel values of the respective stimuli, not their purported units, RMS *pressure*; to get $\log R_\phi$, Poulton presumably divided by 10, but according to Teghtsoonian, Poulton should have divided by 20. Such changes alter the magnitude of $\log R_\phi$, and beg a question: if such confusion over units occurs *within* a modality, how do we cope *across*-modalities? Figure 3 shows 145 points $(\log R_\phi, n)$ from fits of Stevens power function to magnitude estimates of sensation for six sensory modalities, taken from publications *not* associated with S.S. Stevens. Figure 4 displays the points from Fig. 3 that had the largest $\log R_\phi$ for each sensory modality. These points should best fit Teghtsoonian's Law. They do not (solid line); but note a linear data trend (dotted line).

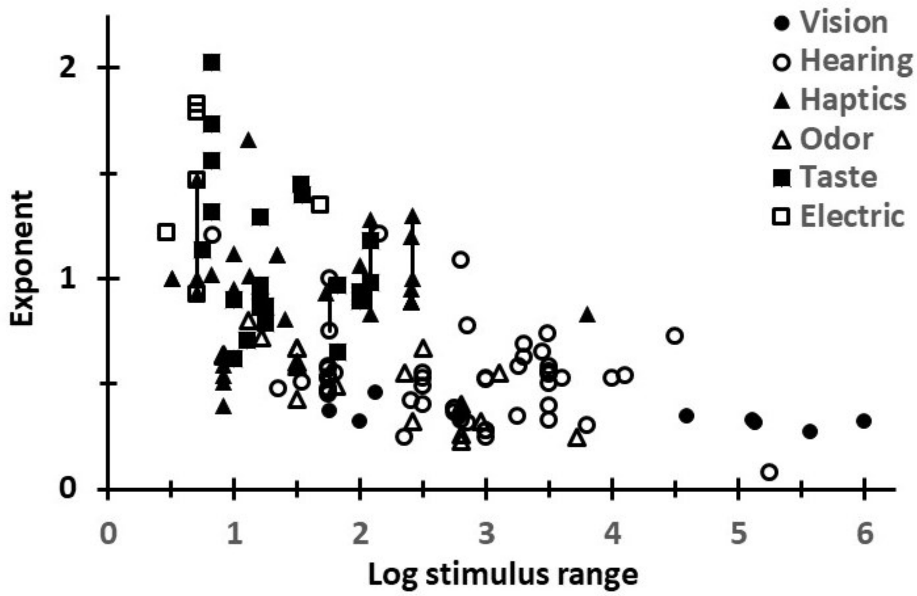


Fig. 3. The sort of data shown in Fig. 2, but this time not by S.S. Stevens and his associates. Vertical bars are exponent ranges. "Haptics" refers to forces exerted upon/by/within the body.

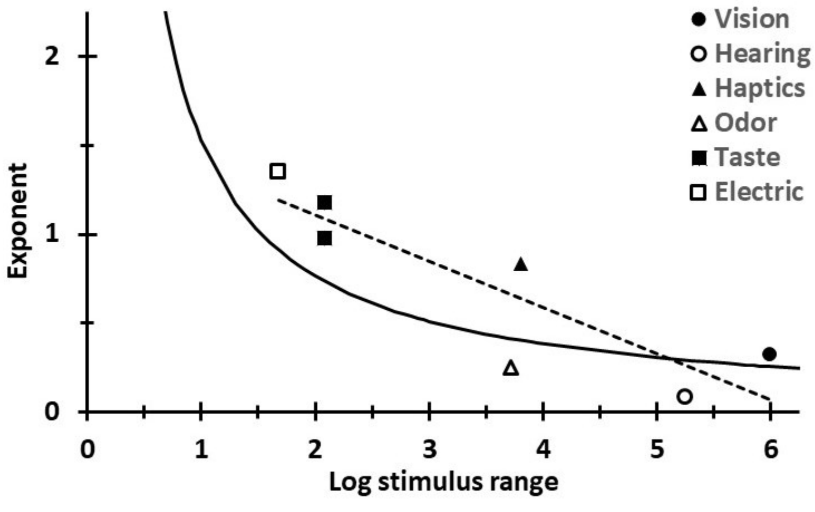


Fig. 4. The rightmost points from Fig. 3 (the two solid squares are a tie). The curve is Teghtsoonian's Law; the fitted dashed line represents a linear trend.

Conclusions

There are no stimulus-magnitude units that are compatible across the senses. For example, no unit describing the magnitude of a tone is equivalent to any unit describing the magnitude of a photon flux. Likewise, ratios cannot be meaningfully compared when units differ across-ratios. Yet Teghtsoonian (1971 – 2018) plotted Stevens power function exponents acquired across-modalities versus the logarithms of the ratios of the highest employed stimulus intensity to the lowest employed stimulus intensity. He found a hyperbolic plot that fitted well to a hyperbolic equation, later deemed Teghtsoonian's Law, that he derived from Stevens' power function. But the correlation can only be artefactual. Indeed, it depends disproportionately upon data-points at intensity-ratio extremes, extremes that change with change of units. The artefactuality of Teghtsoonian's Law becomes clear when it is compared to the Stevens exponents and their respective stimulus-intensity ratios compiled from laboratories not associated with Stevens.

References

Boring, E.G. (1928). Did Fechner measure sensation? *Psychological Review, 35*, 443-445.

Fechner, G.T. (1877). *In Sachen der Psychophysik*. Leipzig, Germany: Breitkopf & Härtel.

Hartmann, W.H. (1998). *Signals, Sound, and Sensation*. New York, NY: Springer-Verlag.

Laming, D. (2009). Weber's Law. In P. Rabbitt (Ed.), *Inside Psychology: A Science Over 50 Years*. New York, NY: Oxford; pp. 179-191.

Myers, A.K. (1982). Psychophysical scaling and scales of physical stimulus measurement. *Psychological Bulletin*, *92*, 203-214.

Poulton, E.C. (1967). Population norms of top sensory magnitudes and S. S. Stevens' exponents. *Perception & Psychophysics*, *2*, 312-316.

Riesz, R.R. (1928). The differential intensity sensitivity of the ear for pure tones. *Physical Review*, *31*, 867-875.

Stevens, S.S. (1962). In pursuit of the sensory law (2nd Klopsteg lecture). Evanston, Ill.: Northwestern University, Technological Institute.

Teghtsoonian, R. (1971). On the exponents in Stevens' law and the constant in Ekman's law. *Psychological Review*, *78*, 71-80.

Teghtsoonian, R. (1973). Range effects in psychophysical scaling and a revision of Stevens' law. *American Journal of Psychology*, *86*, 3-27.

Teghtsoonian, R. (2012). The standard model for perceived magnitude: a framework for (almost) everything known about it. *American Journal of Psychology*, *125*, 165-174.

Teghtsoonian, R. (2018). The concepts of perceived magnitude and dynamic range: what they reveal about the nature of sensory systems. In T. Lachmann & T. Weis (Eds.), *Invariances in Human Information Processing*. New York, NY: Routledge; pp. 1-16.

Teghtsoonian, R. & Teghtsoonian, M. (1978). Range and regression effects in magnitude scaling. *Perception* & *Psychophysics*, *24*, 305-314.

Teghtsoonian, R. & Teghtsoonian, M. (1997). Range of acceptable stimulus intensities: an estimator of dynamic range for intensive perceptual continua. *Perception & Psychophysics*, *59*, 721-728.

Teghtsoonian, R., Teghtsoonian, M., & Karlsson, J.-G. (1981). The limits of perceived magnitude: comparison among individuals and among perceptual continua. *Acta Psychologica*, *49*, 83-94.

Ward, L. (2017). S. S. Stevens' invariant legacy: scale types and the power law. *American Journal of Psychology*, *130*, 401-412.

Weiss, D.J. (1981). The impossible dream of Fechner and Stevens. *Perception*, *10*, 431-434.

HIERARCHICAL PREDICTIVE CODING IN A STATE OF PERCEPTUAL DEPRIVATION: THE OVO-WBPD CHAMBER

Michele Pellegrino[1], Fabio Marson[1], Joseph Glicksohn[2, 3], Mark Elliott[4], and Tal Dotan Ben-Soussan[1]

[1]*Research Institute for Neuroscience, Education and Didactics, Patrizio Paoletti Foundation for Development and Communication, Assisi, Italy*

[2]*Department of Criminology, Bar-Ilan University, Israel*

[3]*The Leslie and Susan Gonda (Goldschmied) Multidisciplinary Brain Research Center, Bar-Ilan University, Israel*

[4]*National University of Ireland, Galway, Ireland*

<m.pellegrino@fondazionepatriziopaoletti.org, research@fondazionepatriziopaoletti.org>

Abstract

The human brain is constantly making predictions about the future and matching them with actual external circumstances in order to provide a stable sense of reality and minimize uncertainty (Bar, 2009). Friston and colleagues (Friston et al., 2006; Friston and Kiebel, 2009) called this process "predictive coding". According to their hypothesis, the cortex employs a top-down prediction algorithm that tries to anticipate incoming sensory stimuli. This algorithm is designed to be dynamic, incorporating not only top-down predictions but also novel stimuli. When predictions and stimuli don't match, a "prediction error" is transmitted to higher-order stages, in which it is used to adjust relevant internal models, that, in turn, will update descending top-down predictions. The outcome of this hierarchical process is a dynamic system that integrates top-down statistical predictions about the environment and actual bottom-up incoming stimuli at multiple levels. Since contemplative practices are aimed cultivating a non-reactive, non-habitual, open attitude towards the full range of experience, a recent study examined the explanatory power of the predictive coding framework also in relation to meditation (Lutz et al., 2019), highlighting the importance of the interoceptive/exteroceptive interplay of information in building accurate models of reality (Jamieson, 2016; Lutz et al., 2019). Yet, so far, the effects of perceptual deprivation on predictive coding have not been examined.

Consequently, we aimed to evaluate, in our ongoing study, the impact of perceptual deprivation and absence of environmental cues on the constant generation and updating of statistical models in the human cortex. We hypothesized that a state of perceptual deprivation could increase the "signal-to-noise" ratio of the incoming stimuli and thus the weight attributed to this by the brain in making prediction about the environment (Friston, 2008).

To this aim, we used a specifically designed room called the OVO-Whole Body Perceptual Deprivation chamber (OVO-WBPD, in the shape of a human sized egg) which has been previously shown to create a state of absorption and alter time perception (Ben-Soussan et al., 2019; Glicksohn et al., 2017).

To test our hypothesis, we employed a previously used method to test different levels of hierarchical predictions (Wacongne et al., 2011; Bekinschtein et al., 2009) in the OVO-WBPD environment. In this auditory task, two levels of predictions are examined. (1) Local: the detection of regularities/irregularities within a sequence of tones, and (2) Global: higher-order level predictions regard the occurring frequency of a given set of stimuli, frequent or infrequent, in a given block. We measured event-related responses known to reflect, within the Predictive Coding framework, the violation of different levels of prediction (the Mismatch Negativity for the Local level and the P3 for the Global level).

To our knowledge, this is the first study to employ a well-documented (Wacongne et al., 2011; Bekinschtein et al., 2009) paradigm testing different levels of prediction errors in a perceptual deprivation environment. Thus, providing new data regarding the importance of exteroceptive and environmental information within the framework of the predictive coding hypothesis.

References

Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1235-1243.

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences, 106(5)*, 1672-1677.

Ben-Soussan, T. D., Mauro, F., Lasaponara, S., Glicksohn, J., Marson, F., & Berkovich-Ohana, A. (2019). Fully immersed: state absorption and electrophysiological effects of the OVO whole-body perceptual deprivation chamber. *Progress in Brain Research, 244*, 165-184.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of physiology-Paris, 100(1-3)*, 70-87.

Friston, K. (2008). Hierarchical models in the brain. *PLoS computational biology, 4(11)*, e1000211.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical transactions of the Royal Society B: Biological sciences, 364(1521)*, 1211-1221.

Glicksohn, J., Berkovich-Ohana, A., Mauro, F., & Ben-Soussan, T. D. (2017). Time perception and the experience of time when immersed in an altered sensory environment. *Frontiers in Human Neuroscience, 11*, 487.

Jamieson, G. A. (2016). A unified theory of hypnosis and meditation states: The interoceptive predictive coding approach. In A. Raz & M. Lifshitz (Eds.), *Hypnosis and meditation: Towards an integrative science of conscious planes* (pp. 313–342). Oxford University Press.

Lutz, A., Mattout, J., & Pagnoni, G. (2019). The epistemic and pragmatic value of non-action: a predictive coding perspective on meditation. *Current opinion in psychology, 28*, 166-171.

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences, 108(51)*, 20754-20759.

THE BLURSDAY PROJECT ON TIME PERCEPTION DURING COVID-19 CONFINEMENT: A GLIMPSE AT THE CANADIAN DATA

Pier-Alexandre Rioux[1], Maximilien Chaumon[2], Antoine Demers[1], Hugo Fitzback-Fortin[1], Sebastian L. Kübel[3, 4], Catherine Lebrun[1], Esteban Mendoza-Duran[1], Luigi Micillo[5], Charles Racine[1], Nicola Thibault[1], Virginie van Wassenhove[6], and Simon Grondin[1]

[1]*École de Psychologie, Université Laval, Québec, QC, Canada*

[2]*Institut du Cerveau, ICM, Inserm U1127, CNRS UMR 7225, Sorbonne Paris - CENIR*

[3]*Institute for Frontier Areas of Psychology and Mental Health, Freiburg, Germany*

[4]*Max Planck Institute for the Study of Crime, Security and Law, Freiburg, Germany*

[5]*Department of General Psychology, University of Padova, Padova, Italy*

[6]*Cognitive Neuroimaging Unit, NeuroSpin, CEA, INSERM, CNRS, Université Paris-Saclay, Gif/Yvette, France*

Abstract

The quarantine associated with the COVID-19 pandemic led to different psychological time-related feelings of strangeness (Grondin et al., 2020). In order to quantify and understand these feelings and the effect of social isolation on temporal information processing, a consortium of 32 time perception researchers from Asia, Europe, and the North and South Americas undertook a global on-line cross-cultural investigation (Chaumon et al. 2022)

In addition to psychophysical experimental tasks, the investigation included measures of stringency and mobility, estimations of the speed of the passage of time and of temporal distances, and questionnaires to determine the chronotype, levels of anxiety and depression, and temporal perspective of participants. The data set, which is available to the community of researchers, can be found at: https://timesocialdistancing.shinyapps.io/Blursday/
Researchers have already begun to extract findings from this dataset. For instance, it has been shown that temporal perspectives were associated with the levels of both anxiety and depression in the first wave of the pandemic in 2020 (Micillo et al., 2022).

The data presented here are from two confinement periods during the COVID-19 pandemic in Canada (spring 2020: $n = 66$; winter 2021: $n = 100$). For each of the confinements, depression levels and the impression of being lonely are negatively associated with a perceived slower flow of time. Moreover, during the second lockdown period, depression is positively associated, whereas the strictness of confinement measures are negatively associated with the impression that future days appear further away. The present study thus contributes to a better understanding the temporal experience of Canadians during the pandemic.

References

Chaumon, M., Rioux, P.-A., Herbst, S., Spiousas, I., Kübel, S., Hiroyasu, E. G., …, & van Wassenhove, V. (2022). The Blursday database as a resource to study subjective temporalities during Covid-19. *Nature Human Behaviour (accepted)*.

Grondin, S., Mendoza, E., & Rioux, P-A. (2020). Pandemic, quarantine and psychological time. *Frontiers in Psychology, section Theoretical and Philosophical Psychology. 11:581036.*

Micillo, L. et al. (2022). Time Perspective predicts levels of Anxiety and Depression during the COVID-19 outbreak: A cross-cultural study. *PLoS ONE (accepted)*.

PERCEPTION OF SPOKEN EMOTIONS IN YOUNG ADULTS WITH INTELLECTUAL DISABILITY

Vered Shakuf[1] and Boaz M. Ben-David[2]

[1]*Department of Communications Disorders, Achva Academic College, Arugot, Israel*

[2]*Communication, Aging and Neuropsychology Lab (CAN lab), School of Psychology, Reichman University (IDC), Herzliya, Israel*

Abstract

In recent years, a lot of efforts have been made to integrate young adults with Intellectual disability (ID) into the community. Effective social interactions are critical for successful community integration. The evidence in the literature suggests that adults with ID have difficulties in identifying emotions in speech, which may in turn negatively affect social interactions. Barriers in emotional speech perception can result from difficulties in understanding the lexical content of the utterance or the prosody (tone of voice). In the current study we adapted the Test for Rating of Emotions in Speech (T-RES), originally designed to test the interplay of semantics and prosody in typical development populations, to test emotional speech perception of adults with ID. The adapted t-res contained 9 sentences neutral in their lexical content that conveyed one of three emotions (anger, happiness or sadness) in the prosodic channel and 9 sentences with neutral prosody that conveyed one of three emotions (anger, happiness or sadness) in the lexical channel. Participants indicated whether they agree that the sentence convey a predefined emotion (anger, happiness or sadness). Our results show that intellectual disability impairs discrimination of spoken emotions, both in the prosodic and semantic channels. As the severity of ID increases, the discriminability of spoken emotions decreases. This hints on the role of primary cognitive abilities in basic identification of spoken emotions.

In recent years, we witness an increase in the efforts to integrate people with Intellectual Disability (ID) into the community. These efforts reflect global trends of disability inclusion and the principle of non-institutionalization (UN Assembly, 2020). Effective interactions between IDs and the general, Typically Developed (TD), population is crucial for successful integration (Forrester-Jones et al., 2006). For example, Migliore and colleagues (2008) found that concerns regarding the social environment are one of the main reasons that most adults with ID work in sheltered workshops. This is despite the fact that they prefer integrated employment and despite government policies that promote supported employment programs. The ability to correctly identify emotions in speech is an essential part of effective and adaptive social interaction. Difficulties in understanding the emotion expressed by the other may lead to communication failure with adverse consequences to well-being and quality of life (Ben-David et al., 2019).

Evidence in the literature suggest that Individuals with ID have difficulties identifying basic emotions expressed via facial expressions (Zaja & Rojahn, 2008; for a review, see Moore, 2001). However, other studies have shown that the use of stimuli with higher ecological reliability, such as videos of dynamic facial expressions or adding context, improved the ability of participants with ID to identify emotions. Moore (2001) suggested that this improvement with more ecological stimuli is evidence that difficulties in performing emotion recognition tasks are not due to specific impairment in emotion perception, but a result of poor information processing capabilities related to low intelligence. This is in contrast to other researchers who claim that people with intellectual disabilities have a specific disability in perceiving emotions, in addition to impaired information processing abilities. For example, Rojahn and his colleagues (1995) found that adults with mild to moderate ID were less accurate in processing facial emotion cues compared to a control group of TD adults and a control group of TD children matched for mental age. However, this was not the case for facial age cues, both the ID adult group and the TD child group were less accurate than the TD adult group. These findings led Rojahn and his colleagues to conclude that impaired ability to identify emotions in facial expressions associated with mental disability cannot be fully explained by mental age.

Many studies have been focused on emotions identification in facial expressions in ID. Less attention has been given to the processing of the auditory channels in this population. In spoken language, emotions can be conveyed via two auditory channels: the lexical channel (semantics, the meaning of the words) and the prosodic channel (intonation of voice, indexical cues). Most of the studies that tested identification of spoken emotions in ID focused on the role of the lexical channel (Gumpel & Wilson, 1996; McKenzie et al., 2001). Recent studies suggest that a major part of the observed difficulty of participants with ID in emotion recognition tasks can be attributed to verbal overload and difficulties in understanding language (Pochon et al., 2017; Wishart et al., 2007). Only a small number of studies have examined IDs ability to identify emotions from prosodic cues, and the results are inconclusive. Hobson, Outson & Lee, (1989) found that adults and adolescents with ID had difficulties (compared to children with typical development matched for mental-verbal age) to match pictures of faces to sounds that express one of six basic emotions. This difference was not found when they were asked to match non-emotional voices and sounds to pictures. These results support the emotion-specificity hypothesis that states that deficits in decoding emotions associated with intellectual disability cannot be fully accounted for by intellectual capabilities. Contrary to the specificity hypothesis, in another study by Hobson and colleagues (1989), participants were asked to label the sounds with a name, and not a picture, adults with ID had difficulties (compared to the TD control group) in both emotional and non-emotional tasks. These differences can be explained by the additional difficulty posed by the verbal demands of the labeling task compared to the picture matching task.

In the current study, we tested the ability of young adults with different levels of intellectual disability to identify the emotional content of the semantics (words) and prosody (tone of speech) of spoken sentences.

Method

Participants

The study included 24 young adults with intellectual disability (Mean age = 24.7). 5 Participants had Mild ID, 10 Mild-Moderate ID and 9 Moderate ID. They were all recruited from a day care center for young adults with intellectual disability.

Tools and Materials

For the purpose of this study we adapted the Hebrew version of Test for Rating of Emotions in Speech (T-RES; Ben-David et al., 2016). The T-RES is a tool designed to test the interplay between the two channels (semantics and prosody) in the perception of emotion in speech. We used the following emotions: Anger, Happiness, Sadness and Neutral. All lexical sentences were equated on main linguistic characteristics (e.g., frequency of usage, sentence length). These sentences were recorded by a native Hebrew Israeli professional radio-drama actress, using the four different prosodies. All sentences were rated as distinctive and good representation of their respective prosodic and lexical categories. To adapt the T-RES to test adults with ID, a group of speech and language pathologists, experienced in working with this population chose a subset of sentences from the original Hebrew T-RES. The adapted T-RES contained 9 sentences neutral in their lexical content that conveyed one of three emotions (anger, happiness or sadness) in the prosodic channel and 9 sentences with neutral prosody that conveyed one of three emotions. We also simplified the instructions and task.

Design and Procedure

Upon arrival, all participants received a detailed explanation regarding the experimental task, followed by a short practice. Each experimental trial began with the presentation of the audio file, followed by the specific instructions presented aurally by a research assistant. Participants were asked if they agree that the speaker conveys a predefined emotion, in three separate emotion blocks (anger, sadness, or happiness). No feedback was provided throughout the task.

Results

The results are presented in Figures 1 and 2. As can be clearly seen. The findings indicate a decrease in the ability to identify emotions in both the semantic and the prosodic channels for all levels of ID tested. This decrease was greater the higher the severity of ID. Young adults with mild ID demonstrate a relatively high discrimination ability. The level of discrimination decreases for mild-moderate level of disability. It seems that this level of discrimination is not preserved in moderate level of ID.

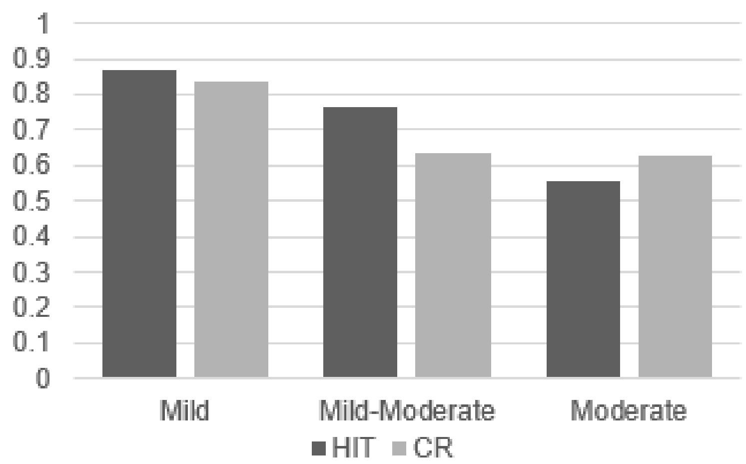


Fig. 1. Proportion of correct responses (Hits and Correct Rejections) to sentences that caries the emotional content in the semantic channel for three levels of intellectual disability: Mild, Mild-Moderate and Moderate

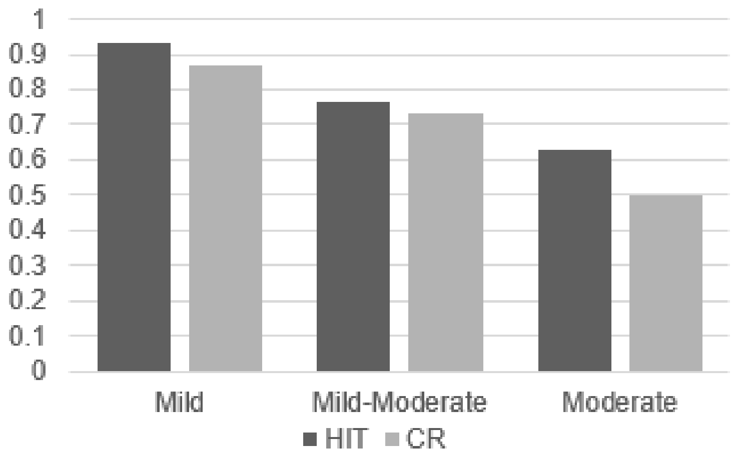


Fig. 2. Proportion of correct responses (Hits and Correct Rejections) to sentences that caries the emotional content in the prosodic channel for three levels of intellectual disability: Mild, Mild-Moderate and Moderate.

Discussion

Identification of emotions in speech is in the core of effective social interaction. To date, only few studies have examined difficulties in understanding the emotions in spoken language of people with intellectual disability and the results are inconclusive. The current study aimed to test identification of emotions in two speech channels: prosody and semantics. Our findings, indicate impaired discrimination of spoken emotions in both speech channels and for all levels of ID tested. However, severity of impairment was directly associated with the severity of the disability. As the severity of ID increases, the discriminability of spoken emotions decreases. These findings highlight the role of primary cognitive abilities in basic identification of spoken emotions and support theories suggesting that difficulties in perceiving emotions stem from a general disability in processing information related to intellectual disability.

In light of the findings showing that among adults with mild ID, the impairment in the ability to identify emotions both in the semantic channel and in the prosodic channel is relatively minor. Further studies should be conducted to test whether the difficulty reported in the literature of adults with mild ID in spoken emotional communication is not necessarily due to difficulty in identification, but rather to the fact that they integrate speech channels differently, compared to the TD population.

Acknowledgements

This study was supported by a grant from the Shalem Foundation [grant number 230].

References

Ben-David, B. M., Gal-Rosenblum, S., van Lieshout, P. H., & Shakuf, V. (2019). Age-Related Differences in the Perception of Emotion in Spoken Language: The Relative Roles of Prosody and Semantics. *Journal of Speech, Language, and Hearing Research*, *62*(4S), 1188–1202.

Ben-David, B. M., Multani, N., Shakuf, V., Rudzicz, F., & van Lieshout, P. H. (2016). Prosody and semantics are separate but not separable channels in the perception of emotional

speech: Test for rating of emotions in speech. *Journal of Speech, Language, and Hearing Research*, *59*(1), 72–89.
Forrester-Jones, R., Carpenter, J., Coolen-Schrijner, P., Cambridge, P., Tate, A., Beecham, J., Hallam, A., Knapp, M., & Wooff, D. (2006). The social networks of people with intellectual disability living in the community 12 years after resettlement from long-stay hospitals. *Journal of Applied Research in Intellectual Disabilities*, *19*(4), 285–295.
Gumpel, T., & Wilson, M. (1996). Application of a Rasch analysis to the examination of the perception of facial affect among persons with mental retardation. *Research in Developmental Disabilities*, *17*(2), 161–171.
Hobson, R., Ouston, J., & Lee, A. (1989). Naming emotion in faces and voices: Abilities and disabilities in autism and mental retardation. *British Journal of Developmental Psychology*, *7*(3), 237–250.
Hobson, R. P., Ouston, J., & Lee, A. (1989). Recognition of emotion by mentally retarded adolescents and young adults. *American Journal on Mental Retardation*.
McKenzie, K., Matheson, E., McKaskie, K., Hamilton, L., & Murray, G. C. (2001). A picture of happiness: Karen McKenzie and colleagues show how to improve communication with people with learning disabilities. *Learning Disability Practice*, *4*(1), 26–29.
Migliore, A., Grossi, T., Mank, D., & Rogan, P. (2008). Why do adults with intellectual disabilities work in sheltered workshops? *Journal of Vocational Rehabilitation*, *28*(1), 29–40.
Moore, D. G. (2001). Reassessing emotion recognition performance in people with mental retardation: A review. *American Journal on Mental Retardation*, *106*(6), 481–502.
Pochon, R., Touchet, C., & Ibernon, L. (2017). Emotion Recognition in Adolescents with Down Syndrome: A Nonverbal Approach. *Brain Sciences*, *7*(6), 55.
Rojahn, J., Rabold, D. E., & Schneider, F. (1995). Emotion specificity in mental retardation. *American Journal on Mental Retardation*.
UN Assembly. (2020). *United Nations disability inclusion strategy (UNDIS)*.
Wishart, J., Cebula, K., Willis, D., & Pitcairn, T. (2007). Understanding of facial expressions of emotion by children with intellectual disabilities of differing aetiology. *Journal of Intellectual Disability Research*, *51*(7), 551–563.

THE DEVELOPMENT AND VALIBATION OF AN ARABIC DIGITS-IN-NOISE TEST

Adnan M. Shehabi[1, 2], Garreth Prendergast[1], Hannah Guest[1], and Christopher J. Plack[1, 3]
[1]*Manchester Centre for Audiology and Deafness, University of Manchester, UK*
[2]*Department of Audiology and Speech Therapy, Birzeit University, Palestine*
[3]*Department of Psychology, Lancaster University, UK*
<adnan.shehabi@postgrad.manchester.ac.uk>

Abstract

The digits-in-noise (DIN) test is a widely used self-administered speech-perception-in-noise hearing screening and auditory research tool that assesses the identification of digits in the presence of background noise. The DIN test has shown high sensitivity and specificity to detecting sensorineural hearing loss. It can be administered using landline phones, portable devices, online, and in the lab. The test has been developed and validated in several world languages including British and American English, French, German, Polish, and Dutch. However, it is still not available in many other world languages.

Aims and Objectives

This study aimed to assess the validity and reliability of newly developed lab-based and online versions of the DIN test in Modern Standard Arabic. This was achieved by determining: (1) the difference between speech-recognition thresholds (SRTs) obtained using the lab-based and online versions of the Arabic DIN test; (2) the difference between the SRTs obtained using the lab-based Arabic DIN and lab-based British English DIN; (3) the difference in the test-retest reliability of SRTs (as measured by the intraclass correlation coefficient, ICC) between the lab-based and online versions of the Arabic DIN; (4) the difference in the test-retest reliability of SRTs (as measured by the ICC) between the lab-based Arabic and lab-based British English versions of the DIN; (5) the difference in the strength of the two correlations for comparisons between: (a) the lab-based Arabic DIN SRT and the standard pure tone audiometric (PTA) threshold average; and (b) the lab-based British English DIN SRT and the standard PTA threshold average; (6) difference in the strength of the two correlations for comparisons between: (a) the lab-based Arabic DIN SRT and the extended high-frequency (EHF) threshold average; and (b) the lab-based British English DIN SRT and the EHF threshold average

Methods

Fifty-two audiometrically normal Arabic- and English-speaking participants aged 18-35 and with no past or current diagnosis of cognitive/memory impairments were recruited. Participants performed two lab-based testing sessions and two at-home online tasks. In the first lab session, participants performed the lab-based Arabic and British English DIN tests, standard pure-tone audiometry at 0.25-8 kHz, extended high-frequency audiometry at 10 and 14 kHz, and otoscopy and tympanometry (to rule out outer and middle-ear pathologies). Upon the completion of the first lab session, participants performed the online Arabic DIN at home twice on two separate days using their own computers/laptops and their choice of headphones/earphones. Participants then completed the second lab-based session, which involved the Arabic and English lab-based DIN tests. The digits and the maskers (speech-shaped noise) for both the lab-based and online versions of the Arabic and English DIN tests were presented diotically and were low-pass filtered at 8 kHz.

Results

(1) The lab-based Arabic DIN SRT (mean: -10.76, SD: 0.92) was similar to, but significantly lower, than the online Arabic DIN SRT (mean: -10.23, SD: 1.144; $t(51) = -2.37$, $p = 0.021$). (2) The lab-based Arabic DIN SRT was significantly higher than the lab-based British English DIN SRT (mean: -11.49, SD: 0.96; $t(51) = 4.2$, $p < 0.0001$). (3) The ICC of the lab-based Arabic DIN SRT (ICC = 0.152) was statistically similar to the ICC of the online Arabic DIN SRT (ICC = 0.328; $p > 0.05$). (4) The ICC of the lab-based Arabic DIN SRT (ICC = 0.152) was statistically similar to the ICC of the lab-based British English DIN SRT (ICC = 0.093; $p > 0.05$). (5) No statistically significant differences were found in the strength of the two correlations for comparisons between (a) the lab-based Arabic DIN SRT and the standard PTA threshold average ($r = 0.068$) and (b) the lab-based British English DIN SRT and the standard PTA threshold average ($r = 0.144$; $p > 0.05$). (6) Similarly, no statistically significant difference in the strength of the two correlations for comparisons between (a) the lab-based Arabic DIN SRT and the EHF threshold average ($r = -0.135$) and (b) the lab-based British English DIN SRT and the EHF threshold average ($r = 0.241$, $p > 0.05$).

Conclusions

The differences in mean SRTs between (1) the lab-based and online versions of the Arabic DIN test and (2) the lab-based Arabic and lab-based British English DIN tests for young audiometrically normal adults were clinically marginal (i.e., < 1.5 dB) with small between-subject variability. The (1) lab-based and online Arabic DIN SRTs and the (2) lab-based Arabic and lab-based British English DIN SRTs both exhibited comparably poor test-retest reliability. The high homogeneity in the standard PTA and EHF thresholds across the participants of the current study may have contributed to the poor correlations between the lab-based Arabic DIN SRT and (1) the PTA and (2) EHF threshold averages. The newly developed lab-based and online versions of the Arabic DIN test may be valid auditory assessment tools. Further research is necessary to establish the relation between Arabic DIN SRTs and the standard PTA and EHF thresholds.

LUMINANCE CONTRAST SENSITIVITY TO INCREMENTS AND DECREMENTS IN THE PRESENCE OF A MOTION MASK WITH AFTERIMAGES: EXPLORING MOTION INDUCED BLINDNESS

Wm Wren Stine[1], Dan Swift[2], Rebecca A. White[1], Sofia C. Lombardo[1], Erika S. Wells[3], John E. Sparrow[4], Joseph A. LaBarre[1], and Andrew J. Kitt[1]
[1]Department of Psychology, University of New Hampshire, Durham, NH, USA
[2]Department of Behavioral Sciences, University of Michigan-Dearborn, Dearborn, MI, USA
[3]Department of Psychological and Brain Sciences, Boston University, Boston, MA, USA
[4]Department of Life Sciences, University of New Hampshire-Manchester, Manchester, NH, USA
<bill.stine@unh.edu>

Abstract

Motion-Induced Blindness (MIB) describes the disappearance of a salient visual stimulus in the presence of a motion mask. We measured the inhibition created by the motion mask as a luminance contrast threshold. Increment and decrement targets, presumably stimulating ON-channels and OFF-channels, respectively, exhibited mask-induced inhibition when in the presence of an increment or decrement mask, with greater inhibition when the targets and masks were of the same contrast polarity. Further, with greater mask-induced inhibition, afterimages lost influence over thresholds.

MacKay (1960) showed that the perception of a retinally-stabilized image of a 10 min of arc diameter wire, which normally would not fade when stabilized, vanished when an unstabilized image was visually scanned. Grindley and Townsend (1965) then discovered that when a target is presented to one eye and a moving mask to the other eye, the target will appear to disappear, which they described as movement masking. Ramachandran and Gregory (1991) and Spillmann and Kurtenbach (1992) found that a uniform patch in the periphery would appear to vanish if surrounded by dynamic random-dot noise, with the noise appearing to fill in the patch (see Kawabe & Miura, 2007; Wallis & Arnold, 2008, as well). Ramachandran and Gregory (1991) describe perceptual filling-in as forming an artificial scotoma. Bonneh et al. (2001) asked subjects to fixate a point foveally and attend to one or more targets (small yellow dots) in the periphery while a mask of discreet elements (small blue '+', for example) moved coherently such that any element intersecting a target moved behind that target. The targets again appeared to disappear. Bonneh et al. (2001) coined the term Motion Induced Blindness (MIB) to describe this effect.

There is evidence suggesting that perceptual filling-in and MIB shared common mechanisms (Hsu, Yeh, and Kramer, 2004; 2006), as might be expected given that perceived motion would seem to be related to motion energy normalized by flicker energy (i.e., motion contrast: Georgeson & Scott-Samuel, 1999; Rainville et al., 2005; Rainville et al., 2002).

Caetta et al. (2007) suggest that the invisibility of the targets *per se* involve changes in sensitivity to the target and shifts in detection criteria, which may be modeled as response gain changes, as measured by brightness matching, and contrast gain changes, as measured using contrast detection thresholds (Gorea & Caetta, 2009). Total disappearance time is *positively* related to target contrast, negatively related to target size and speed, and positively related to mask contrast, dot density, and speed, as well as exhibiting Gestalt grouping effects for both the target and the mask (Bonneh et al., 2001; Graf et al., 2002, Experiment 2; Mitroff & Scholl, 2005; Shibata et al., 2010). Stine et al. (2017) demonstrated that decrement masks induce increment target disappearance more quickly than increment masks, and, overall, increment targets disappear more quickly than decrement targets, perhaps due to target and mask contrast valence stimulating ON- and OFF-channels (Dolan, & Schiller, 1994; Schiller, 1992; Schiller et al., 1986; Zaghloul et al., 2003). Libedinsky et al. (2009) suggest that the effect of the mask is to increase the likelihood of disappearance rather than to decrease the visibility of the target. Interestingly, Dieter, Tadin, and Pearson (2015) demonstrated that the MIB process continues even when continuous flash suppression (Tsuchiya & Koch, 2005) is used to remove the process from visual awareness.

So, contrast valence effects (Stine et al., 2017; White et al., 2020) on disappearance time are consistent with what one might expect from a simple model where the observer's motion mask response alters contrast gain (cf., Caetta et al., 2007; Gorea & Caetta, 2009). In this experiment we wished to replicate and extend the work of White et al. (2020) across all combinations of stimulus contrast valence. Building on our previous work, we again used a stimulus with four peripheral dots, or inducers, and a motion mask consisting of 64 incoherently moving dots, but where both the inducers and mask may have positive (brighter than the background), negative (darker than the background), or zero (the same as the background) contrast valence. As before, four seconds into the trial, the four peripheral inducers were physically removed and, after a variable delay, a single dot, or target, was briefly flashed in one of the original four inducer locations. The target varied in contrast valence, with either positive or negative contrast. The observer then reported in which location the target appeared, giving a four alternative forced choice paradigm.

Measuring target contrast detection thresholds, we predicted that a non-zero valence would have an impact relative to zero contrast for both the mask, which thus would induce MIB, and the inducer, due to the afterimage created, and that there would be differences between positive and negative contrasts such that within 'channel' effects would be stronger than between channel effects. For the target contrast, we also predicted a difference between positive and negative contrasts with within-channel stronger than between-channel. For the inducer to target inter-stimulus intervals, we predicted that the biggest effect would be at the shortest interval, again reflecting the afterimage, with decreasing effects toward longer intervals.

One might expect that, as the mask-induced inhibition waned, target contrast threshold would decrease with the time interval between the offset of the inducers and the onset of the target. But this effect was missing in White et al.'s (2020) data. Finally, replicating previous work, decrements should have a lower threshold than increments (Dolan, & Schiller, 1994; Schiller, 1992; Schiller et al., 1986; Zaghloul et al., 2003).

Methods

Participants

Three females and one male, over the age of 18 with normal or corrected to normal vision, participated in the study. The study was approved by the University of New Hampshire Institutional Review Board.

Apparatus and Stimuli

A Dell Dimension E521 computer running *Vision Works* (Swift et al., 1997) in *Windows XP* drove a Image Systems M21L-H4101 monitor with a 120 Hz refresh rate. The monitor uses a monochrome P46 ultra-short persistence phosphor (yellow-green; CIE x = 0.427, y = 0.543), presenting 800 x 600 pixels with a pixel pitch of 120 dots per inch. Gray scale was rendered using a Vision Research Graphics Gray-Scale Expander VW16 to provide 15-bit linearized depth. A 50 cd/m^2 background was continuously present. Positive contrast stimuli (increments) were 100 cd/m^2 and negative contrast stimuli (decrements) were 0 cd/m^2. The 21" flat-screen monitor was viewed at distance 1 m while using a chin rest.

Four independent variables were manipulated. Three of these variables involved the contrast valence of the stimuli, which were 16 arc-minute circular targets on a uniformly gray background. They were presented in a temporal Gaussian window (100 ms SD) for 200 ms. The mask and inducers were increments (positive contrast), decrements (negative contrast), or absent (zero contrast), while the targets were either increments or decrements. The inducers were four deg of retinal angle from the fixation dot. The mask moved incoherently at 4 deg of retinal angle per s. The fourth independent variable was the inter-stimulus interval between the offset of the inducers and the onset of the target, varying across six delays. All four subjects experienced each of the 3 x 3 x 2 x 6 = 108 conditions at least 20 times presented in random order.

Procedure

Sitting in a darkened room with their head stabilized by a chin rest 1 m from the monitor, each participant viewed the central fixation dot and adapted to the background luminance of the screen for five min. Following adaptation, each trial lasted 20 s during which the motion mask was continuously visible, except in the absent condition while the four inducers were present just during the first four seconds of the trial, again except in the absent condition. At delays of 0.5 s, 3.5 s, 6.5 s, 9.5 s, 12.5 s, and 15.5 s after the initial 4 s, the target was briefly flashed (250 ms) in one of the original four inducer locations. After each flash an auditory beep indicated to the participant a brief response interval during which he or she could report the location of the target's flash using a keyboard. The participant received feedback after each of the six responses during the trial. The participant was instructed to maintain fixation on the centrally located fixation dot. A 22 s adaptation period followed each trial.

As a function of the time interval between the offset of the inducer and the target onset, inducer stimulus valence (with contrasts of +100%, -100%, or 0%), motion mask contrast valence (+100%, -100%, or 0%) and target contrast valence (increment or decrement), the contrast of the target was varied following a weighted up-down adaptive psychophysical procedure (Smith, 1961; Kaernbach, 1991) in order to converge onto a 0.625 probability of a hit. Inducer stimulus contrast valence, mask contrast valence, and the target contrast valence were randomly varied across trials while the inducer offset to target onset varied within trials, as mentioned previously.

Results

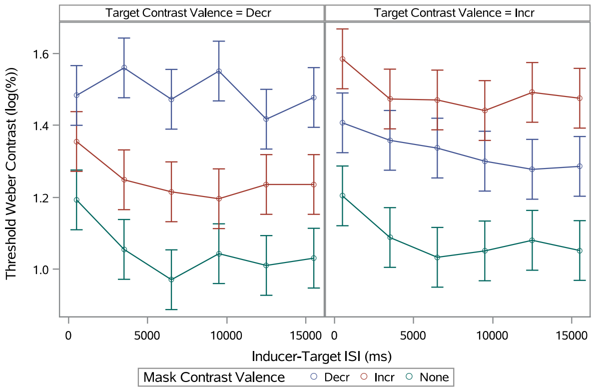
We fit a normal cumulative density function to the probability of a hit as a function of target log Weber percent contrast weighted by the number of presentations for each condition, which varied according to the weighted up-down procedure. The means and standard deviations from those fits served as our raw dependent variables. While analyzing these two variables independently, we chose a test type I error rate of $0.0253 = 1 - (1 - 0.05)^{1/2}$, using the Šidák (1967) inequality (see Kirk, 2013, p. 183), to keep the familywise type I error rate equal to 0.05. Only the analysis of the psychometric function means will be reported here.

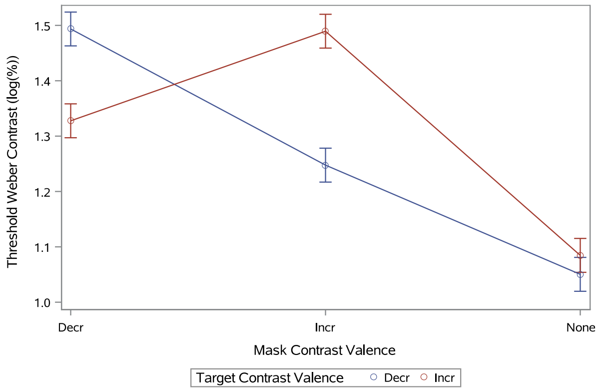
We conducted a repeated measures randomized-block factorial analysis of variance (RBF-3326; Kirk, 2013, Ch. 10.5-10.9) using *SAS* v9.4. Our approach was to partition each main effect sum of squares into orthogonal one degree of freedom *a priori* contrasts. For inducer and mask contrast valence, we predicted that a non-zero valence would have an impact relative to zero contrast, and that there would be differences between the positive and negative contrasts.

For the target contrast, we also predicted a difference between positive and negative contrasts. For the inducer to target inter-stimulus intervals, we predicted that the biggest effect would be at the shortest interval, with decreasing effects toward larger intervals (i.e., a Helmert pattern comparing the interval of interest to the average of the remaining longer intervals). All contrast-contrast interactions of order two through four, which were orthogonal to one another and to the main-effect contrasts, were also tested. All contrasts were tested using error terms calculated specifically for each contrast as sphericity was not tenable.

Preliminary analyses showed the mean log Weber percent contrasts to be normally distributed and a Tukey test for non-additivity for the design was significant ($F(1, 320) = 8.83$, $p = 0.0032$). As well, there were differences among the participants ($F(3, 321) = 57.62$, $p < 0.0001$, partial $\rho_I = 0.3439$).

The presence of the motion mask interacted with the contrast valence of the target when comparing 500 ms inducer to target ISI – a contrast-by-contrast-by-contrast interaction ($F(1, 3) = 28.62$, $p = 0.0128$, partial $\widehat{\omega}^2 = 0.0601$). When the motion mask was absent (green in Figure 1), thresholds were low and there was an after-image effect lasting fewer than 3500 ms. If the motion mask was the same valence as the target, thresholds were maximally elevated. When the motion mask and target were of opposite valence, thresholds were above that without the mask but lower than when the valences were the same as one another. Further, as thresholds increased, the difference between an ISI of 500 ms and the average of the greater ISIs decreased.





Averaging over other effects, the presence of the inducer increased thresholds at 500 ms ISI relative to longer ISIs ($F(1, 3) = 26.21$, $p = 0.0144$, partial $\widehat{\omega}^2 = 0.0551$). As well, when the contrast valence of the mask and that of the target matched one another, target contrast thresholds were increased ($F(1, 3) = 17.63$, $p = 0.0247$, partial $\widehat{\omega}^2 = 0.0371$).

Discussion

Replicating Stine et al. (2017) and White et al. (2020), contrast valence influences MIB. If increment contrasts tap ON-channels and decrement contrasts OFF-channels (Dolan, & Schiller, 1994; Schiller, 1992; Schiller et al., 1986; Zaghloul et al., 2003), motion masks are more effective within channels than between channels.

Thresholds as a function of inducer-target ISI flattened as thresholds increased, so the presumed afterimage influence on threshold evaporated with increased inhibition due to the motion mask over and above the evident rise in contrast detection threshold. As stated previously by White et al. (2020, p. 59) “Presumably, residual inhibition of the visibility of the inducers from the motion mask (i.e., motion induced blindness) reduced the target’s contrast gain (i.e., raised the target’s contrast threshold).” And again, we saw no effect of inducer-target

ISI; once the inducers have been inhibited, that inhibition can be maintained by the motion mask.

These findings are consistent with those of LaBarre and Stine (2019), Stine at al. (2017), and White et al. (2020) that the increment – decrement distinction influences motion induced blindness.

Acknowledgement

We would like to thank our respective institutions for their support.

References

Bonneh, Y. S., Cooperman, A., & Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature, 411*, 797-801.

Caetta, F., Gorea, A., & Bonneh, Y. S., (2007). Sensory and decisional factors in motion-induced blindness. *Journal of Vision, 7*(7), 1–12. doi:10.1167/7.7.4

Dieter, K. C., Tadin, D., & Pearson, J. (2015). Motion-induced blindness continues outside visual awareness and without attention. *Scientific Reports, 5*, 11841

Dolan, R. P. & Schiller, P. H. (1994). Effects of ON channel blockade with 2-amino-4-phosphonobutyrate (APB) on brightness and contrast perception in monkeys. *Visual Neuroscience, 11*, 23–32.

Georgeson, M. A., & Scott-Samuel, N. E. (1999). Motion contrast: A new metric for direction discrimination. *Vision Research, 39*, 4393–4402.

Grindley, G. C. & Townsend, V. (1965). Binocular masking induced by a moving object. *Quarterly Journal of Experimental Psychology, 17*, 97-109. doi:10.1080/17470216508416418

Gorea, A., & Caetta, F., (2009). Adaptation and prolonged inhibition as a main cause of motion-induced blindness. *Journal of Vision, 9,* 1-17. doi:10.1167/9.6.16

Graf, E. W., Adams, W. J., & Lages, M. (2002). Modulating motion-induced blindness with depth ordering and surface completion. *Vision Research, 42*, 2731-2735.

Hsu, L.C., Yeh, S.L., & Kramer, P. (2004). Linking motion-induced blindness to perceptual filling-in. *Vision Research, 44*, 2857-2866.

Hsu, L.C., Yeh, S.L., & Kramer, P. (2006). A common mechanism for perceptual filling-in and motion-induced blindness. *Vision Research, 46*, 1973–1981.

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics, 49*, 227-229

Kawabe, T., & Miura, K. (2007). Subjective disappearance of a target by flickering flankers. *Vision Research, 47*, 913–918.

Kirk, R. E. (2013). *Experimental design: Procedures for the social sciences* (4^{th} ed.). Washington, DC: Sage.

LaBarre, J. A., & Stine, W. W. (2019). The effects of motion adaptation and contrast polarity on motion-induced blindness. In N. Du Bois, S. Arndt, E. B. Özsoy, S. Bayraktar, E. Gülbetekin, & M. A. Elliott (Eds.). *Fechner day 2019: Proceedings of the 35th annual meeting of the international society for psychophysics* (pp. 44-51). Antalya, Turkey: International Society for Psychophysics.

Libedinsky, C., Savage, T., & Livingstone, M. (2009). Perceptual and physiological evidence for a role for early visual areas in motion-induced blindness. *Journal of Vision, 9*, 1–10

MacKay, D. M. (1960). Monocular "rivalry" between stabilized and unstabilized retinal images. *Nature, 185*, 834.

Mitroff, S. R., & Scholl, B. J. (2005). Forming and updating object representations without awareness: Evidence from motion-induced blindness. *Vision Research, 45*, 961–967.
Rainville, S. J. M., Makous, W. L., & Scott-Samuel, N. E. (2005). Opponent-motion mechanisms are self-normalizing. *Vision Research, 45*, 1115–1127.
Rainville, S. J. M., Scott-Samuel, N. E., & Makous, W. L. (2002). The spatial properties of opponent-motion normalization. *Vision Research, 42*, 1727–1738.
Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature, 350*, 699-702.
Schiller, P. H. (1992). The ON and OFF channels of the visual system. *Trends in Neurosciences, 15*, 86-92
Schiller, P. H., Sandell, J. H., & Maunsell, J. H. R. (1986). Functions of the ON and OFF channels of the visual system. *Nature, 322*, 824-825
Shibata, M., Kawachi, Y., & Gyoba, J. (2010). Combined effects of perceptual grouping cues on object representation: Evidence from motion-induced blindness. *Attention, Perception, & Psychophysics, 72*, 387–397.
Šidák, Z. K. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association, 62,* 626–633.
Smith, K. J. E. (1961). Stimulus programming in psychophysics. *Psychometrika, 26*, 27-33
Spillmann, L., & Kurtenbach, A. (1992). Dynamic noise backgrounds facilitate target fading. *Vision Research, 32*, 1941-1946.
Stine, W. W., Levesque, P. A., Lusignan, M. E., & Kitt, A. J. (2017). Motion induced blindness using increments and decrements of luminance. *Proceedings of the Latvian Academy of Sciences. Section B. Natural, Exact, and Applied Sciences, 71*, 372–379.
Swift, D., Panish, S., & Hippensteel, B. (1997). The use of *VisionWorks* in visual psychophysical research. *Spatial Vision, 10*, 471-477.
Wallis, T. S. A., & Arnold, D. H. (2008). Motion-induced blindness is not tuned to retinal speed. *Journal of Vision, 8,* 1–7. doi:10.1167/8.2.11.
White, R. A., Swift, D., Lombardo, S. C., Wells, E. S., Sparrow, J. E., Oostendorp, A., LaBarre, J. A., Kitt, A. J., & Stine, W. W. (2020). Motion induced blindness: Luminance contrast sensitivity to increments and decrements in the presence of a motion mask. In J. R. Schoenherr, T. Hubbard, W. W. Stine, & C. Leth-Steensen (Eds.), *Proceedings of the 36th annual meeting of the international society for psychophysics: Fechner day 2020* (pp. 56-59). International Society for Psychophysics.
Zaghloul, K. A., Boahen, K., & Demb, J. B. (2003). Different circuits for ON and OFF retinal ganglion cells cause different contrast sensitivities. *The Journal of Neuroscience, 23*(7), 2645-2654.

REVISITATION TO THE EFFECTS OF PRECEDING SOUNDS ON TIME-SHRINKING AND AUDITORY TEMPORAL ASSIMILATION

Hiroshige Takeichi[1] and Emi Hasuo[2]

[1]*Open Systems Information Science Team, Advanced Data Science Project (ADSP), RIKEN Information R&D and Strategy Headquarters (R-IH), RIKEN, Yokohama, 230-0045 Japan*

[2]*JSPS research fellow at Hiroshima University (Center for Brain, Mind, and Kansei Sciences Research), Hiroshima, 734-8553 Japan*

<takeichi@riken.jp, emi8513@gmail.com>

Abstract

Time-shrinking is an illusion in time perception: an empty time interval (*t2*) is perceived to be shorter than its actual duration when it is preceded by an adjoining and physically shorter empty time interval ($t1$) ($-\ 50\ ms \leq t1 - t2\ \leq 80\ ms$). The illusion can be demonstrated by having another comparison stimulus, following but separated from the test stimulus (i.e., *t*1 and *t*2), adjusted to be subjectively equal to *t*2 in duration. The comparison stimulus at the point of subjective equality (PSE) will be shorter than the physical limen of *t*2, or PSE of *t*2 presented alone (without $t1$). This apparent reduction of subjective duration ("shrinking") of *t*2 can be also demonstrated through judgment on perceived equality between $t1$ and *t*2. The *t*2 when $t1$ and *t*2 are perceived to be equal will be longer than the physical limen of *t*2 at *t*1 = *t*2, (i.e., *t*1). This phenomenon is called auditory temporal assimilation. We hypothesized that this illusion stems from ambiguity in decomposing the temporal pattern into tempo and rhythm. For example, if *t2* is longer than *t1*, it may be interpreted either as an isochronous tone train in decreasing tempo, or a non-isochronous tone train at a constant tempo, being inherently ambiguous if they are presented in isolation. In order to examine the feasibility of this hypothesis, we made an informal (n=1) observation of the three-tone stimulus with preceding tone-trains that either accelerate or decelerate in tempo. The preceding tone-trains comprised of four tone bursts of 440 Hz. The individual intervals were either increased or decreased in a geometric series of duration by 5% per step, i.e., [82%, 86%, 91%, 95%] or [123%, 117%, 111%, 105%] of *t*1+*t*2, such that the train smoothly connects and provides a natural context to the three tone bursts of 880 Hz. The tone bursts had a duration of 20 ms and included a rise and a fall time of 5 ms. The total duration *t*1+*t*2 was varied in 4 steps: 240, 360, 480, and 600 ms. For each combination of the accelerating and decelerating contexts and the 4 total durations, the matching task (time shrinking) and the judgment task (auditory temporal assimilation) were performed, using a 32-series random staircase method (divided and conducted in two 3-hour lasting sessions) in which the position of the second (middle) tone of the three tones was varied to change *t*2 in each trial. The *t*1 duration thus changed contingently with *t*2 because *t*1+*t*2 was held constant. As a result, while the illusion persisted with the accelerating context in both tasks, it was lost or inverted with the decelerating context. Although the result indicates an effect of the context of changing tempo, the sign of the effect was the opposite of the naïve interpretation above: physically longer *t*2 was perceived to compose an isochronous rhythm when the context implied a physically shorter *t*2 with the accelerating tempo. The leading tones with changing tempo do not seem to merely provide cues to the expected marker position. Instead, the results may suggest independent but interacting processing of event timing (displacement) and changing tempo (derivatives). The illusion, at any rate, provides a tool to clarify as to how predictive-coding of multi-scale time perception characterizes human auditory perception and cognition particularly in music.

MAGNETOENCEPHALOGRAPHY OF PROCESSING MOSAIC AND CHECKERBOARD SPEECH

Hiroshige Takeichi[1], Kazuo Ueda[2], Takako Mitsudo[3], Alexandra Wolf[4, 3], and Yoji Hirano[3]
[1]*Open Systems Information Science Team, Advanced Data Science Project (ADSP), RIKEN Information R&D and Strategy Headquarters (R-IH), RIKEN, Yokohama, 230-0045, Japan*
[2]*Department of Acoustic Design/Research Center for Applied Perceptual Science/Research and Development Center for Five-Sense Devices, Kyushu University, Fukuoka 815-8540, Japan*
[3]*Department of Neuropsychiatry, Graduate School of Medical Sciences, Kyushu University, Fukuoka, 812-8582, Japan*
[4]*Special Postdoctoral Researcher (SPDR) at RIKEN Center for Advanced Intelligence Project, Cognitive Behavioral Assistive Technology Team, Chuo City, Tokyo 103-0027 Japan*
<takeichi@riken.jp, ueda@design.kyushu-u.ac.jp>

Abstract

While accurate and precise dynamics or temporal development appears to be critical to speech intelligibility, it is in fact insensitive to local time reversal or local averaging to some extent. This led to a hypothesis that speech perception is processed in two temporal windows. Recent studies additionally clarified that frequency resolution interacts with these temporal windows, suggesting a spectrotemporally constrained grouping in speech perception. The purpose of the present study is to search for the neural signature of the time windows and the spectrotemporal grouping in order to clarify the mechanisms of human speech perception. Magnetoencephalograms (MEG) were acquired while the participant was passively listening to the original and degraded speech stimulus (mosaic speech: noise-vocoded with time-frequency patches, and checkerboard speech: interrupted with time and frequency segments), using a 306 channel whole head MEG system (Neuromag, Elekta/MEGIN). In Experiment 1 (n=17), eight types of the stimuli were randomly presented: (1) the original, (2) noise-vocoded (3) mosaic speech with 20 frequency bands and 40-ms segment duration (20-40), (4) 20-160, (5) 20-320, (6) 4-40 (the 4 bands identified through factor analysis of speech), (7) 4-40 (the 4 bands defined by combining the critical bands), and 2-40. In Experiment 2 (n=10), seventeen types were randomly presented: interrupted speech stimuli with 16 frequency bands and 20-, 80-, 160- and 320-ms segment duration, checkerboard speech stimuli with 12 combinations of 4, 8 or 16 frequency bands and 20-, 80-, 160- or 320-ms segment duration, and filtered speech stimuli with the 16 frequency bands and modulation with the 320-ms segments as the baseline control. After downsampling from 1000 Hz to 250 Hz, high-pass filtering with a cut-off at 0.1 Hz and band-reject filtering at 60 Hz and 120 Hz were performed. Then, the data were screened with artifact rejection by using fast-independent component analysis and kurtosis criteria. The ratio of the mean power during the listening period to the mean power during the pre-stimulus one-second baseline period was calculated for each stimulus separately for four frequency bands: theta (4 – 8 Hz), alpha (8 – 13 Hz), beta (13 – 30 Hz), and gamma (30 – 50 Hz). As a result, the alpha power > the beta power and the gamma power for the original speech, theta > beta and theta > gamma for the noise vocoded speech, and alpha > gamma for the 20-bands 320-ms mosaic, and alpha > beta for the 2-bands 40-ms mosaic speech. Whereas the increased alpha power for the mosaic speech was likely to reflect drowsiness induced by unintelligible stimulus, the increased theta power for the noise-vocoded stimuli, which was absent for the other speech stimuli, may reflect speech processing.

OLFACTORY PERCEPTUAL EVALUATION OF NEW-BORN INFANTS: PRELIMINARY RESULTS OF A NEW SMELL SCALE

Rosana M Tristão[1], Elaine R Neiva[2], Kassandra S F Costa[3], Laiane Medeiros Ribeiro[3], Geraldo M Fernandes[1], Karina N Costa[1], Jan Spilski[4], and Thomas Lachmann[4, 5]
[1]*Faculty of Medicine and University Hospital, University of Brasília, Brasília, Brazil*
[2]*Institute of Psychology, University of Brasília, Brasília, Brazil*
[3]*Faculty of Health and University Hospital, University of Brasília, Brasília, Brazil*
[4]*Center for Cognitive Science, University of Kaiserslautern, Germany*
[5]*Facultad de Lenguas y Educación, Universidad Nebrija, Spain*

Abstract

Olfactory loss in adults is one of the initial and most frequent acute clinical manifestations of the SARS-COV-2 infection and olfactory loss has been one major symptom. The few studies that have evaluated this alteration in the paediatric range have shown that children have less olfactory sensory loss than adults. The goal of the present study is to develop and validate a behavioural evaluative scale of olfactory perception for infants born from women who tested positive for COVID-19 during pregnancy compared to no at-risk new-borns. This is an observational comparative analytical cohort study of 96 new-borns exposed and unexposed to COVID-19 during pregnancy. The data collection is an experimental procedure exploring odours of the maternal breastmilk, vanilla (sweet), coffee (acid/bitter) and distilled water (neutral). There was no difference between the groups (control and sick) in terms of the mothers' gestation time. The severity of maternal illness had effect over olfactory responses from infants compared to the control group. There was no effect over the responses from developing and clinical variables such as maternal stress, Apgar, height, weight and brain circumference and neurological clinical examination as well as for gender. The judge's concordance for the responses achieved different levels and it was included responses with moderate to high correlation values. The infants were able to detect and discriminate among odours and control and at-risk groups differed in response patterns. Head orientation and mouth movements were at the most powerful categories of behaviours to detection and discrimination thresholds and were considered to compose the final version of the Smell Scale for Newborn Infants.

Maternal SARS-CoV-2 infection does not expose the foetus and new-born only to the effects caused directly by the virus, but also a variety of indirect effects. The consequences of the maternal and foetal inflammatory response, with the production of potentially cytotoxic cytokines, as well as the effect of the use of antiviral medications have not been studied to date. Another important aspect concerns the risk of contamination of the new-born during or shortly after birth. Practices such as delayed cord clamping and skin-to-skin contact between mothers and new-borns are not universally recommended and evidence on the risk of contagion during breastfeeding is still limited (Ashokka et al., 2020; Yu et al., 2020). Among the symptoms provided by carriers of the new coronavirus, is an olfactory sensory alteration. Olfactory loss in adults is one of the earliest and most frequent acute clinical manifestations of SARS-CoV-2 infection. It is feasible to hypothesize the involvement of the foetus' olfactory bulb during intrauterine life as one of the indelible pathophysiological manifestations to the clinical diagnosis of COVID-19 with neurosensory olfactory deficit in foetuses and new-borns affected by intrauterine infection. Based on this evidence, the following research question was raised: do new-born children of women infected with SARS-CoV-2 during pregnancy have olfactory sensory changes? Hence, this work is based on the scientific literature on the sensory-perceptual development of smell in foetuses and new-born babies (André et al., 2018; Bartocci et al., 2000; Bingham et al., 2007; Marlier et al., 1998; Moura et al., 2014; Sarnat et al., 2017; Schaal, 2000; Schaal et al., 1998) and aims to replicate methodological procedures of studies such as Bartocci et al. (Bartocci et al., 2000), adapting this to the hospital context and expanding it as categories of olfactory stimulus for the implementation of a measure of discrimination sensitivity. The main goal of the study is to assess the olfactory sensory perception of new-born children of women infected with SARS-CoV-2 during pregnancy. The specific objectives are to develop a behavioural evaluation scale of olfactory perception in infants; to validate this scale in healthy new-borns; and to compare the responses with that of new-born children of women infected with COVID-19 during pregnancy.

Methodology

This observational comparative analytical study analysed the behavioural responses of 96 new-borns up to 14 days old, whose mothers tested positive compared to asymptomatic mothers who tested negative, based on a test that detects genetic material of the virus using a laboratory technique called polymerase chain reaction (PCR) for SARS-CoV-2 during pregnancy. The data collection followed an experimental procedure from Bartocci et al (Bartocci et al., 2000) that explores odours of the maternal breastmilk, vanilla (sweet), and distilled water (neutral) and coffee (acid/bitterness) was added. Each test epoch consists of 30 s of baseline definition followed by 30 s of smell exposure with a two-minute interval for washout effect (Figure 1).

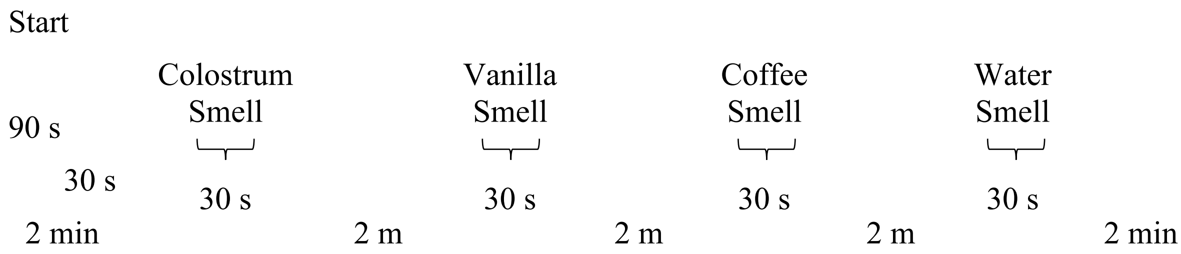


Figure 1. Timeline of odour exposure during one session adapted from Bartocci et al (Bartocci et al., 2000).

The study was approved by the Research Ethics Committee from the University of Brasilia School of Medicine (http://www.fm.unb.br/cep-fm - CAAE 32359620.0.0000.5558) and was registered in the Brazilian Registry of Clinical Trials (ReBec) under number RBR-65qxs2. The complete protocol can be accessed at ReBec “Effects of COVID-19 on pregnancy, childbirth, puerperium, neonatal period and child development” [cited 2020 Sep 17]. Available from: http:// ensaiosclinicos.gov.br/rg/ RBR-65qxs2 /).

Six judges, three Brazilians and three Germans analysed the sessions videos frame by frame to survey categories of behavioural responses to the four different olfactory stimuli. The rating levels considered two response categories “yes” or “no”, for present and absent reaction to the smell, also registered the duration of response, intensity, and valence. Concordance among judges for the infants’ responses was compared using Nonparametric Correlations analyses. Demographic and clinical data were collected. Factor Analysis was performed with analyses of the main components for the construction of the instrument and analysis of the sample data. A mixed design was run, with treatment (colostrum, vanilla, coffee, water) as within-subject variable and Covid (covid at-risk vs. control babies) as between-subject variable (group variable). Continuous variables were described as mean and standard deviation or median and interquartile range (IQR), as appropriate. Categorical variables were described as frequencies and percentages. The analyses were run to determine the perception thresholds, sensitivity parameters, and ROC curves (Receiver Operating Characteristic). In addition to fixed effects models, random effects models were estimated and adjusted for possible confounders. Statistical significance was considered when the *p* value is less than .05.

Preliminary Results

A range of 36 responses were collected from the video analyses and were distributed among four behavioural categories: Head-orientation, Mouth movement, Facial movements, and General body movement. The judges' level of concordance achieved different levels of correlational and significance values depending on the response category. Responses that did not achieve a moderate level of concordance among judges were excluded. In general, the accepted concordance level was $r>.60$. The more robust correlations were found to responses from Head Orientation (e.g.: head approaching or eyes turning to the odorant to milk or head avoiding the odorant to water and coffee, $r \geq .80$, $p \leq .030$) and Mouth Movements (e.g.: tongue protrusion $r \geq .80$, $p \leq 0.00$; sucking $r \geq .90$, $p \leq 0.001$; smiling, lips parting or lips puckering $r \geq .8$, $p \leq 0.001$) categories. There was no difference between the groups (control and sick) in terms of the mothers' gestation time. There was no effect over the responses from developing and clinical variables such as maternal stress, Apgar, height, weight and brain circumference and neurological clinical examination as well as for gender. However, the severity of maternal illness had effect over olfactory responses from infants compared to the control group varying to the type of response category ($p \leq .050$).

Discussion

Hitherto, no study has explored the clinical and behavioral aspects of odours responses of new-born infants born from mothers who were infected by the SARS-COV-2 virus (COVID) during pregnancy, hence, at risk for being infected in utero. The infants, control and at-risk groups, were able to detect and discriminate among the four categories of odours and they differed in various levels of response. Among all behavioural categories, the avoidance or approaching movements and mouth movements achieved the highest values of concordance among judges being candidates to compose the final scale.

This was a preliminary set of analyses for this study and alternate models will be applied to explore and refine the results, though at the first glance these findings have already shown the impact of pregnancy infection over the newborn infants' odour response. Pregnant women, as a special group, can be infected with SARS-CoV-2, with a possible consequent infection of their foetuses and new-borns. SARS-COV-2 infection can cause an immune overreaction that is manifested by the excessive activation of immune cells and the production of a large amount of interferon and cytokines that can affect foetal development and increase the risk of neurological diseases in the neonatal period (Dong et al., 2020). One of the proposed mechanisms for altering smell and taste related to COVID-19 is the ability of SARS-CoV-2 to bind to ACE2 in the nasal and oral mucosa. Among people who tested positive for COVID-19, olfactory sensory sensation was significantly less impaired in children than in adults ($p = .00014$). The significant difference in olfactory sensory impairment between children and adults, and particularly between younger children and middle-aged adults, is in line with the finding that the expression of ACE2 in the nasal epithelium and in the oral cavity is more intense among adults, corroborating the possibility that the distribution and expression of ACE2 in the oral cavity and nasal epithelium may contribute to differences in sensory impairment (22). Little is known about foetal impairment in mothers infected with SARS-CoV-2. By evaluating the olfactory sensory perception of new-borns of women infected with COVID-19 during pregnancy, we can quantify such involvement.

References

André, V., Henry, S., Lemasson, A., Hausberger, M., & Durier, V. (2018). The human newborn's umwelt: Unexplored pathways and perspectives. *Psychonomic Bulletin & Review*, *25*(1), 350–369. https://doi.org/10.3758/s13423-017-1293-9

Ashokka, B., Loh, M.-H., Tan, C. H., Su, L. L., Young, B. E., Lye, D. C., Biswas, A., E Illanes, S., & Choolani, M. (2020). Care of the Pregnant Woman with COVID-19 in Labor and Delivery: Anesthesia, Emergency cesarean delivery, Differential diagnosis in the acutely ill parturient, Care of the newborn, and Protection of the healthcare personnel. *American Journal of Obstetrics and Gynecology*, S0002937820304300. https://doi.org/10.1016/j.ajog.2020.04.005

Bartocci, M., Winberg, J., Ruggiero, C., Bergqvist, L. L., Serra, G., & Lagercrantz, H. (2000). Activation of Olfactory Cortex in Newborn Infants After Odor Stimulation: A Functional Near-Infrared Spectroscopy Study. *Pediatric Research*, *48*(1), 18–23. https://doi.org/10.1203/00006450-200007000-00006

Bingham, P. M., Churchill, D., & Ashikaga, T. (2007). Breast milk odor via olfactometer for tube-fed, premature infants. *Behavior Research Methods*, *39*(3), 630–634. https://doi.org/10.3758/BF03193035

Dong, L., Tian, J., He, S., Zhu, C., Wang, J., Liu, C., & Yang, J. (2020). Possible Vertical Transmission of SARS-CoV-2 From an Infected Mother to Her Newborn. *JAMA*. https://doi.org/10.1001/jama.2020.4621

Marlier, L., Schaal, B., & Soussignan, R. (1998). Neonatal responsiveness to the odor of amniotic and lacteal fluids: A test of perinatal chemosensory continuity. *Child Development*, *69*(3), 611–623.

Moura, R. G. F., Cunha, D. A., Gomes, A. C. de L. G., & Silva, H. J. da. (2014). Quantitative instruments used to assess children's sense of smell: A review article. *CoDAS*, *26*(1), 96–101. https://doi.org/10.1590/s2317-17822014000100014

Sarnat, H. B., Flores-Sarnat, L., & Wei, X.-C. (2017). Olfactory Development, Part 1: Function, From Fetal Perception to Adult Wine-Tasting. *Journal of Child Neurology*, *32*(6), 566–578. https://doi.org/10.1177/0883073817690867

Schaal, B. (2000). Human Foetuses Learn Odours from their Pregnant Mother's Diet. *Chemical Senses*, *25*(6), 729–737. https://doi.org/10.1093/chemse/25.6.729

Schaal, B., Marlier, L., & Soussignan, R. (1998). Olfactory function in the human fetus: Evidence from selective neonatal responsiveness to the odor of amniotic fluid. *Behavioral Neuroscience*, *112*(6), 1438–1449. https://doi.org/10.1037/0735-7044.112.6.1438

Yu, N., Li, W., Kang, Q., Xiong, Z., Wang, S., Lin, X., Liu, Y., Xiao, J., Liu, H., Deng, D., Chen, S., Zeng, W., Feng, L., & Wu, J. (2020). Clinical features and obstetric and neonatal outcomes of pregnant patients with COVID-19 in Wuhan, China: A retrospective, single-centre, descriptive study. *The Lancet Infectious Diseases*, *20*(5), 559–564. https://doi.org/10.1016/S1473-3099(20)30176-6

AUDITORY GROUPING FACILITATES UNDERSTANDING INTERRUPTED MOSAIC SPEECH STIMULI

Kazuo Ueda[1], Hiroshige Takeichi[2], and Kohei Wakamiya[3]
[1]*Department of Acoustic Design, Faculty of Design/Research Center for Applied Perceptual Science/Research and Development Center for Five-Sense Devices, Kyushu University, Fukuoka 815-8540, Japan*
[2]*Open Systems Information Science Team, Advanced Data Science Project (ADSP), RIKEN Information R&D and Strategy Headquarters (R-IH), RIKEN, Kanagawa 230-0045, Japan*
[3]*Department of Acoustic Design, Faculty of Design, Kyushu University, Fukuoka 815-8540, Japan*
<ueda@design.kyushu-u.ac.jp, takeichi@riken.jp, wakamiya@design.kyushu-u.ac.jp>

Abstract

The intelligibility of interrupted speech stimuli has been known to be almost perfect when segment duration is shorter than 80 ms, which means that the interrupted segments are perceptually organized into a coherent stream under this condition. However, why listeners can successfully group the interrupted segments into a coherent stream has been largely unknown. Here we show that the intelligibility for mosaic speech, in which original speech was segmented in frequency and time, and noise-vocoded with the average power in each unit, was largely reduced by periodical interruption. At the same time, the intelligibility could be recovered by promoting auditory grouping of the interrupted segments with stretching the segments up to 40 ms and reducing the gaps, provided that the number of frequency bands was enough (≧4) and the original segment duration was equal to or less than 40 ms. The interruption was devastating for mosaic speech stimuli, very likely because a poor grouping cue, which resulted from the deprivation of periodicity and temporal fine structure with mosaicking, prevented successful auditory grouping for the interrupted segments. These results suggest that a grouping cue should play an important role in the perception of normal speech under adverse conditions.

Miller and Licklider (1950) provided systematic data on perception of interrupted speech stimuli with a list of phonemically balanced monosyllabic speech stimuli. The study showed that, when the percentage of the time that the speech was left on was 50% and the segment duration (for both on and off) was gradually extended, the speech intelligibility was around 80% until the segment duration became about 80 ms, and the intelligibility declined to about 50% at the 500-ms segment duration. Thus, the interrupted speech segments are perceptually organized into a coherent stream when the segment duration is less than 80 ms. *Mosaic speech* (Fig. 1b; Nakajima et al., 2018; Eguchi et al., 2022) is a degraded speech, both in frequency and temporal resolutions. First, an original speech signal is divided into several frequency bands and time segments. Then, the average power within each unit is reflected to a stepwise amplitude envelope in each frequency band, which modulates a band-noise source. The modulated band-noises are summed across frequency to produce mosaic speech. Here we report that the intelligibility for spectrotemporally degraded speech stimuli, i.e., mosaic speech stimuli, was severely deteriorated with interruption, and that the intelligibility was recovered when we modified the stimuli in a way to promote auditory grouping.

Ueda et al. (2021) provided preliminary results which indicated that the interruption introduced in mosaic speech stimuli (Fig.1c) resulted in a devastating effect: the intelligibility decreased to less than 10% in most of the cases. There were two possible hypotheses to explain the results: the *insufficient speech fragments* hypothesis claimed that insufficient amounts of speech segments caused the deterioration, whereas, the *failed auditory grouping* hypothesis claimed that failure in grouping between adjacent segments caused the deterioration. However, the possibility to stretch the *mosaicked* segment duration had never been explored to examine these hypotheses (Fig.1d-e; cf. Santi et al., 2020). Furthermore, stretching may lead to deteriorated intelligibility (the *reduced temporal resolution* hypothesis).

Therefore, the purpose of the current investigation was twofold: (1) to confirm the preliminary results by Ueda et al. (2021), and (2) to verify the three hypotheses by examining the effect of stretching the mosaicked segment duration on intelligibility, keeping the total duration unchanged by reducing the silent gap duration.

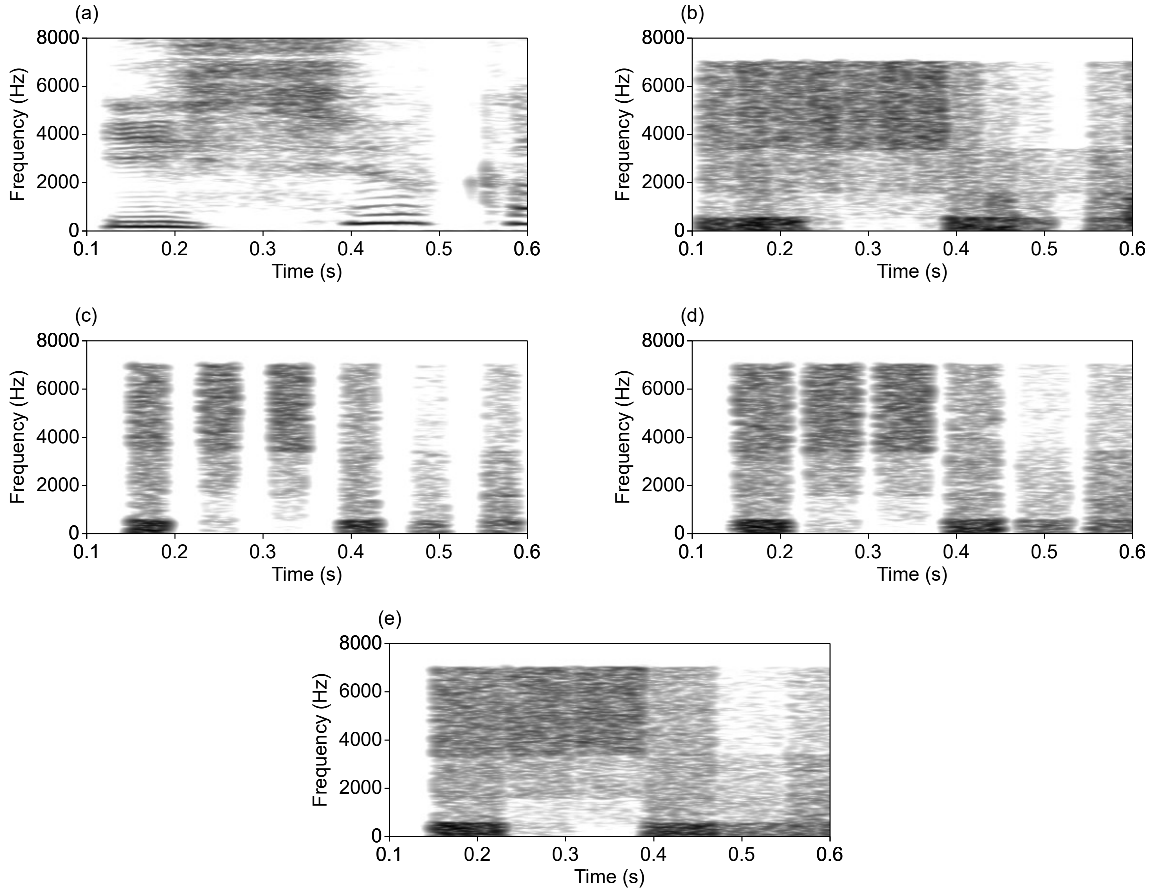


Fig. 1. Examples of narrowband spectrograms for the stimuli, produced from the same fragment of an original spoken Japanese sentence by a female talker from NTT-AT, "Phonemically Balanced 1000-Sentence Speech Database." (a) The original speech sample, (b) mosaic speech with 4 frequency bands (passbands of 50-570, 570-1600, 1600-3400, 3400-7000 Hz) and the 40-ms segment duration including 5-ms root-of-raised-cosine ramps in amplitude (4 band-noises were modulated with stepwise functions reflecting the averaged power in each frequency-time unit), (c) interrupted mosaic speech, i.e., every other segment was replaced with a silent gap, (d) interrupted and stretched mosaic speech in which each segment was stretched in time by a factor of 1.5 and the silent gaps were shrunk by a factor of 0.5, and (e) interrupted and stretched mosaic speech in which each segment was stretched in time by a factor of 2.0 and the silent gaps were removed.

Method

Participants

A total of 12 paid Japanese listeners (ages 20-23) with normal hearing (tested with RION AA-56) participated in the experiment. They were all Japanese native listeners with normal hearing. The research was conducted with prior approval of the Ethics Committee of Kyushu University (approval ID: 70). All the participants gave informed consent in compliance with the protocol approved by the Committee. All methods employed were in accordance with the guidelines provided by the Japanese Psychological Association.

Conditions and Stimuli

Five steps of the number of frequency bands (2, 4, 8, 16, and 20), three steps of original segment duration (20, 40, and 80 ms), and three steps of stretching ratios (1.0, 1.5, and 2.0) were combined to construct experimental conditions. In addition, the original (unprocessed) speech stimuli were employed as control stimuli. The 20 frequency bands were based on critical bands, and the 2-16 frequency bands were based on the 4 frequency bands derived from the factor analysis of speech in eight languages/dialects (Ueda and Nakajima, 2017).

A total of 230 Japanese sentences spoken by a professional male talker were extracted from the "Phonemically Balanced 1000-Sentence Speech Database" (NTT-AT; 44100-Hz sampling, 16-bit linear quantization). The speech samples were mosaicked (Nakajima et al., 2018), interrupted, and stretched. The silent gap duration was adjusted to keep the original total duration.

Procedure

The stimuli were presented to participants diotically through headphones (Beyerdynamic DT 990 PRO) in a double-walled sound-attenuated booth (Music cabin SD3). Participants were instructed to write down exactly what they heard with hiragana or katakana (sets of symbols that are used to represent Japanese morae; a mora is a syllable-like unit in Japanese). They were instructed to write down the morae that they could immediately recognize, and not to fill blanks afterwards from the context.

Results

Figures 2 shows the experimental results. The control condition with the original speech stimuli resulted in almost perfect performance (96%). The obvious trend in Figure 2 for experimental conditions was that intelligibility (measured as the percentages of mora accuracy) decreased with increasing original segment duration and decreasing number of frequency bands, except for the 2-band condition which always resulted in the floor performance. The effect of stretching is most obvious in the data for the 20-ms original segment duration. For example, for the 20-band stimuli, stretching increased mora accuracy from 63% with the stretching ratio of 1.0 (no stretching; Fig. 2a), through 85% with the stretching ratio of 1.5 (Fig. 2b), to 89% with the stretching ratio of 2.0 (Fig. 2c), which was very close to the performance in the control condition. Although the performance generally went down as the number of frequency bands decreased, the improvements in performance with stretching were obvious, except for the 2-band stimuli. The same tendency was still observable to smaller extent for the 40-ms original segment duration, whereas all performance went down to the floor for the 80-ms original segment duration.

To sum up, all three variables manipulated in the experiment, i.e., original segment duration, number of frequency bands, and stretching ratio, were effective in altering intelligibility, except for the 2-band stimuli. These observations were supported by the analysis using a generalized linear mixed model (GLMM) with a logistic linking function as implemented in an add-in for JMP Pro (SAS Institute Inc., 2021).

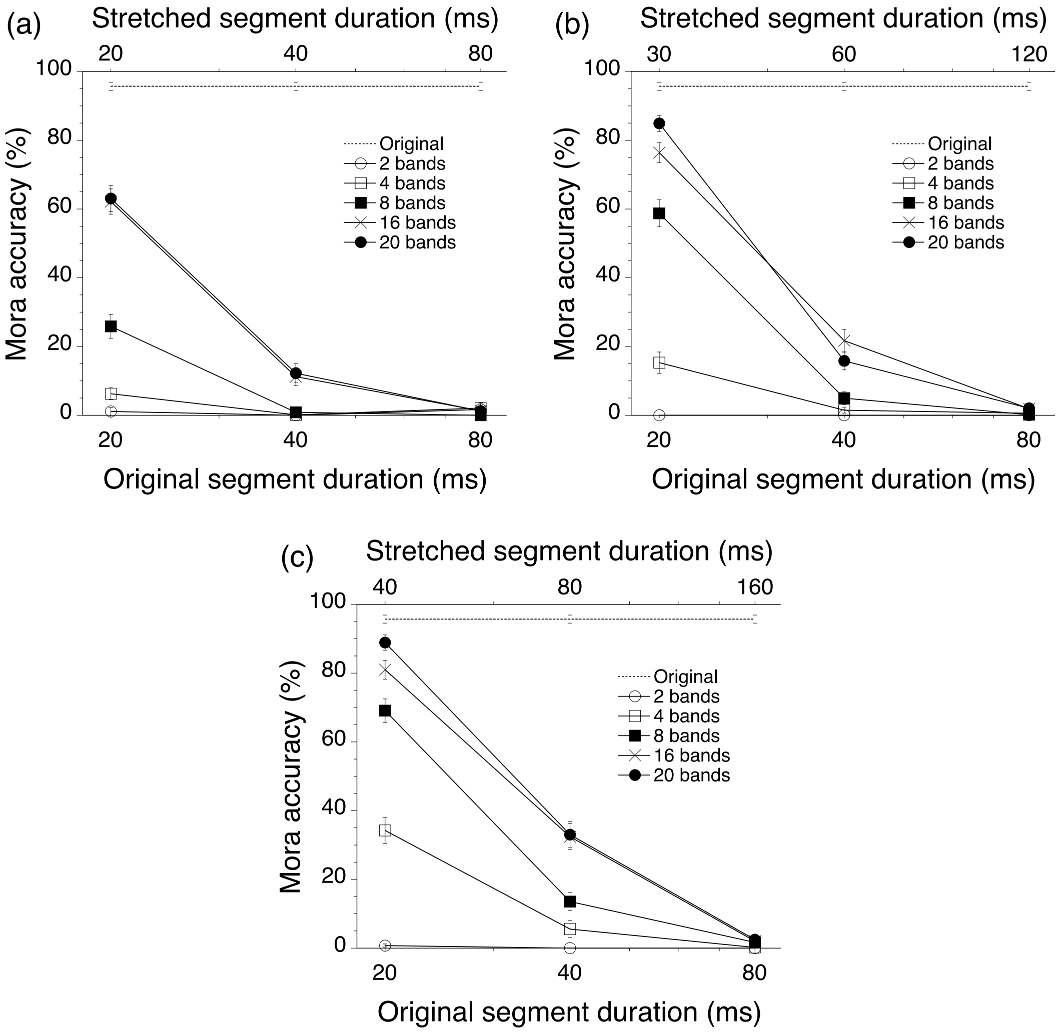


Figure 2. Mora accuracy in percentages as a function of stretching ratio, original segment duration, and number of frequency bands. Stretching ratio: (a) 1.0, i.e., no stretching, every other segment was replaced with a silent gap of the same duration; (b) 1.5, e.g., each original 20-ms segment was stretched to 30 ms, whereas each silent gap was shortened to 10 ms; (c) 2.0, e.g., each original 20-ms segment was stretched to 40 ms and connected to the next segment without any gap. The data for the original stimuli are presented in all panels as a reference. Error bars, standard error of the mean (SEM).

Discussion

The present results successfully replicated the preliminary results by Ueda et al. (2021). Further, the results supported the failed auditory grouping hypothesis and rejected the insufficient speech fragments hypothesis by showing improvements in intelligibility with stretching, because stretching the mosaic segments should facilitate auditory grouping of the segments, but should not increase the number of any speech fragments. At the same time, the reduced temporal resolution hypothesis was also rejected. It was suggested that previously unavailable speech cues became available when the mosaicked segments were stretched and auditory grouping was promoted.

The periodicity and temporal fine structure in normal speech mainly provide a grouping cue (Apoux and Healy, 2013; Apoux et al., 2013) to combine the interrupted segments perceptually (Clarke et al., 2016). Other notable aspects in the perception of interrupted speech stimuli would be the influence of spectral coherence and regularity in interruption. Based on a series of systematic experiments in which checkerboard noise (Howard-Jones and Rosen, 1993) was used, Fogerty et al. (2018, 2020) argued that periodic glimpsing (Buss et al., 2004; Hall et al., 2008; Ozmeral et al., 2012) led to better performance than random spectrotemporal glimpsing, and that spectral coherence in maskers tended to result in better performance compared with temporal coherence. Furthermore, the periodic interruption for the vocoded speech stimuli can be seen as an interfering modulation which is superimposed over an envelope modulation in each frequency band conveying speech information. The situation can be regarded as modulation masking taking place (Houtgast and Steeneken, 1973; Takahashi and Bacon, 1992), and stretching the mosaic segments may reduce the prominence of this interfering envelope components, reducing modulation masking. One may conjecture that this reduction in modulation masking accounts the improvement in the intelligibility for the stretched mosaic speech stimuli, instead of auditory grouping. On the other hand, within the scope of the current investigation, the situation of reducing modulation masking always promotes auditory grouping as well. Considering that wider experimental results can be explained by auditory grouping than the reduction of modulation masking, we are inclined to presume that auditory grouping would be the key element. The arguments above can be generalized into understanding why the intelligibility of interrupted mosaic speech is so devastating and why the intelligibility for normal speech is so robust against periodic interruption.

In conclusion, the number of frequency bands should be equal to or more than four and the original segment duration should be equal to or less than 40 ms, if stretching the mosaicked segments was effective in improving intelligibility. These limits seem to stem from the general limits in speech perception.

Acknowledgements

This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant No. JP19H00630. The authors would like to thank Yoshitaka Nakajima for providing J language software routines, Hikaru Eguchi for programming in LiveCode, Rikito Takahashi for running the experiment, and Koutaro Munechika for valuable comments on the draft.

References

Apoux, F., & Healy, E. W. (2013). A Glimpsing Account of the Role of Temporal Fine Structure Information in Speech Recognition. In B. C. J. Moore, R. D. Patterson, I. M. Winter, R. P. Carlyon, & H. E. Gockel (Eds.), *Basic Aspects of Hearing: Physiology and Perception* (Vol. 787, pp. 119-126).

Apoux, F., Yoho, S. E., Youngdahl, C. L., & Healy, E. W. (2013). Role and relative contribution of temporal envelope and fine structure cues in sentence recognition by normal-hearing listeners. *The Journal of the Acoustical Society of America*, 134(3), 2205-2212.

Buss, E., Hall, J. W., & Grose, J. H. (2004). Spectral integration of synchronous and asynchronous cues to consonant identification. *The Journal of the Acoustical Society of America*, 115(5), 2278-2285. doi:10.1121/1.1691035

Clarke, J., Başkent, D., & Gaudrain, E. (2016). Pitch and spectral resolution: A systematic comparison of bottom-up cues for top-down repair of degraded speech. *The Journal of the Acoustical Society of America*, 139(1), 395-405. doi:10.1121/1.4939962

Eguchi, H., Ueda, K., Remijn, G. B., Nakajima, Y., & Takeichi, H. (2022). The common limitations in auditory temporal processing for Mandarin Chinese and Japanese. Scientific Reports, 12(1), 3002. doi:10.1038/s41598-022-06925-x

Fogerty, D., Carter, B. L., & Healy, E. W. (2018). Glimpsing speech in temporally and spectro-temporally modulated noise. *The Journal of the Acoustical Society of America*, 143(5), 3047-3057. doi:10.1121/1.5038266

Fogerty, D., Sevich, V. A., & Healy, E. W. (2020). Spectro-temporal glimpsing of speech in noise: Regularity and coherence of masking patterns reduces uncertainty and increases intelligibility. *The Journal of the Acoustical Society of America*, 148(3), 1552-1566. doi:10.1121/10.0001971

Hall, J. W., Buss, E., & Grose, J. H. (2008). Spectral integration of speech bands in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 124(2), 1105-1115. doi:10.1121/1.2940582

Houtgast, T., & Steeneken, H. J. M. (1973). The modulation transfer function in room acoustics as a predictor of speech intelligibility. *Acoustica*, 28, 66-73.

Howard-Jones, P. A., & Rosen, S. (1993). Uncomodulated glimpsing in "checkerboard" noise. *The Journal of the Acoustical Society of America*, 93(5), 2915-2922. doi:10.1121/1.405811

Miller, G. A., & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *The Journal of the Acoustical Society of America*, 22(2), 167-173. doi:10.1121/1.1906584

Nakajima, Y., Matsuda, M., Ueda, K., & Remijn, G. B. (2018). Temporal resolution needed for auditory communication: Measurement with mosaic speech. *Frontiers in Human Neuroscience*, 12(149), 1-8. doi:10.3389/fnhum.2018.00149

Ozmeral, E. J., Buss, E., & Hall, J. W. (2012). Asynchronous glimpsing of speech: Spread of masking and task set-size. *The Journal of the Acoustical Society of America*, 132(2), 1152-1164. doi:10.1121/1.4730976

Santi, Nakajima, Y., Ueda, K., & Remijn, G. B. (2020). Intelligibility of English Mosaic Speech: Comparison between Native and Non-Native Speakers of English. *Applied Sciences*, 10(19), 1-13. doi:10.3390/app10196920

SAS Institute Inc. (2021). JMP Pro (version 16.2.0). Cary, NC: SAS Institute Inc.

Takahashi, G., A., & Bacon, S., P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *Journal of Speech, Language, and Hearing Research*, 35(6), 1410-1421. doi:10.1044/jshr.3506.1410

Ueda, K., Kawakami, R., & Takeichi, H. (2021). Checkerboard speech vs interrupted speech: Effects of spectrotemporal segmentation on intelligibility. *JASA Express Letters*, 1(7), 075204. doi:10.1121/10.0005600

Ueda, K., & Nakajima, Y. (2017). An acoustic key to eight languages/dialects: Factor analyses of critical-band-filtered speech. *Scientific Reports*, 7(42468), 1--4. doi:10.1038/srep42468

A MESOLIMBIC VALENCE COMPETITION MODEL EXPLAINS ACUTE PAIN DYNAMICS

Andrew D. Vigotsky, Rami Jabakhanji, Marwan N. Baliki, and A. Vania Apkarian
Departments of Biomedical Engineering & Statistics, Northwestern University, Evanston, IL
Center for Translational Pain Research, Northwestern University, Chicago, IL
Departments of Neuroscience, Anesthesiology, and Physical Medicine & Rehabilitation, Northwestern University, Chicago, IL
Shirley Ryan AbilityLab, Chicago, IL

Abstract

Temporal fluctuations in the intensity of a noxious stimulus produce complex and often unintuitive changes in pain. For instance, offset analgesia is the disproportionately large decrease in pain following a slight decrease in the intensity of a noxious stimulus,[1] dissociating the noxious stimulus from pain via a history-dependent mechanism. These complex dynamics have been captured using phenomenological models derived from system identification;[2,3] however, such models are a black box and provide little insight into the neural mechanisms that give rise to acute pain's complex dynamics. Since a circuit's behavior is governed by its architecture, there is much to learn from generative, circuit motif-based models—a correctly specified model is sufficient to account for the observed data. Here we introduce a circuit-inspired model of acute pain dynamics.

Our model is based on the novel hypothesis that the mesolimbic system's positive and negative reward mechanisms compete with one another. For example, when the intensity of a noxious stimulus is reduced, it is associated with a positive reward since the organism has presumably found a path to safety. Our hypothesis posits that such a positive reward would inhibit the circuits responsible for encoding negative valence, effectively preventing the experience of pain. The circuit model reflecting our hypothesis can be reduced to a pair of coupled, first-order differential equations containing between 5 and 8 interpretable parameters in total, depending on the assumptions one is willing to make about the symmetry of the positive-negative valence competition. We assessed our model's ability to capture the pain dynamics associated with onset hyperalgesia, offset analgesia, and stepwise increases with sequential offset analgesia.[4]

Our mesolimbic valence competition model captured the salient dynamics of all three tasks, lending preliminary support to our novel hypothesis. Importantly, our model also produced interpretable parameters linked to our mechanistic hypothesis of offset analgesia and other dynamical acute pain phenomena.

References

Grill, J. D. & Coghill, R. C. (2002). Transient analgesia evoked by noxious stimulus offset. *Journal of Neurophysiology, 87*, 2205-2208, doi:10.1152/jn.00730.2001.

Cecchi, G. A. et al. (2012). Predictive dynamics of human pain perception. *PLoS Computational Biology, 8*, e1002719, doi:10.1371/journal.pcbi.1002719.

Petre, B. et al. (2017). A central mechanism enhances pain perception of noxious thermal stimulus changes. *Scientific Reports, 7*, 3894, doi:10.1038/s41598-017-04009-9.

Derbyshire, S. W. G., Long, V. J. E. & Asplund, C. L. (2019). Stepwise increasing sequential offsets cannot be used to deliver high thermal intensities with little or no perception of pain. *Journal of Neurophysiology, 122*, 729-736, doi:10.1152/jn.00007.2019.

CHARACTERISTICS OF THREE-DIMENSIONAL OBJECT RECOGNITION IN VIRTUAL REALITY ENVIRONMENTS

Natsuko Wasaki[1, 2] and Tatsuto Takeuchi[1]
[1]*Graduate School of Integrated Arts and Social Sciences, Japan Women's University, 2-8-1 Mejirodai, Bunkyo-ku, Tokyo 112-8681, Japan*
[2]*Research Fellow of Japan Society for the Promotion of Science*
<natsukowasaki@gmail.com>

Abstract

To understand the underlying mechanisms of three-dimensional (3D) object recognition, we examined the visual search performance of an object and its parts. We found a high sensitivity toward the vertical parts compared to the oblique or horizontal parts of objects presented in virtual reality environments. As the visual search performance between the main object (cube) and its vertical parts was almost the same, we depended on the vertical parts of the object for 3D object recognition. We speculate that the dependency on the vertical parts stems from the fact that our daily lives are rich in vertical components represented in spatiotopic reference frames and provide more information than horizontal components.

Understanding the way in which people recognize and identify objects in their daily lives is a fundamental question in the field of vision science. Any object comprises different parts. Therefore, it is reasonable to hypothesize that the recognition of any object depends mainly on the way in which individuals perceive its parts. The structural description theory of three-dimensional (3D) object recognition asserts that the parts are described as generalized cylinders (Marr & Nishihara, 1978) or "geons" (Biederman, 1987). Simple objects, such as cubes, can be divided into several parts (Humphreys & Donnelly, 2000). The purpose of this study was to investigate the mechanism behind the recognition of objects and their parts. Through the use of a visual search paradigm, we examined whether some parts of an object are processed at a higher degree compared to other parts, or whether all parts of an object are processed equally when it comes to object recognition. If people are more sensitive to some specific types of parts compared to other types of parts in an object, we can conclude that object recognition depends mainly on the detection of such parts.

In daily life, people observe 3D objects from many points of view by moving their heads or bodies, thereby allowing for objects to be represented in the object-centered (or spatiotopic) reference frames of their visual systems. In structural description theory, different parts of an object are represented in the spatiotopic reference frame, thereby making object recognition viewpoint-independent. Therefore, in this study, we evaluated the visual search performance of an object and its parts in a virtual reality (VR) environment in which the participants could freely move their heads and bodies. Visual stimuli were randomly placed in the virtual environment and presented to the observers through a head-mounted display.

Method

Participants

A total of 82 participants (average age = 23.7 years) participated in the experiments. The number of participants required for this study was determined using a priori power analyses based on our previous experiments. All the participants had normal or corrected-to-normal vision.

Apparatus

In the VR environment, visual stimuli were presented through a head-mounted display (HMD) (VIVE Pro Eye, HTC Corp) at a resolution of 1440 × 1600 pixels, with a refresh rate of 90 Hz controlled using a Dell Precision 5820 (Dell Inc.) personal computer. The participants used a hand-held controller to provide their responses throughout the experiments. The real-time 3D development platform, Unity (Version 2019.3, Unity Technologies), was used to generate and present the stimuli, control the experimental sequences, and collect data.

Visual stimuli

We used elongated thin cylinders as the element stimuli based on Marr & Nishihara's generalized cylinder. The participants were given the task of finding the target embedded in the distractors (1900 single cylinders) within the virtual environment. As shown in Figure 1, we prepared various types of target stimuli. The logic behind defining the target involved generating a basic object, after which several cylinders were deleted from the generated basic object to show its constituent parts. Both the basic objects and their constituent parts served as targets (Figure 1). As shown in Figure 1A, the basic object was a "cube" constructed using 12 cylinders. We prepared two types of orientations for this cube (0° and 45°). The other parts were generated through the systematic removal of several cylinders from the cube. By removing eight cylinders, we generated "parallel" parts with three different orientations (0°, 45°, and 90°) or "square" parts with two different orientations (0° and 45°). The parts were vertical when the orientation was 0°, and they were horizontal when the orientation was 90°. Through a similar procedure, we generated other basic objects and their parts, as shown in Figures 1B and 1C. As shown in Figure 1B, the basic object was also a cube comprising 12 cylinders. To examine the preciseness of the vertical parts, we slightly tilted each cylinder ("rods"). As shown in Figure 1C, the basic object was generated through the combination of three cubes.

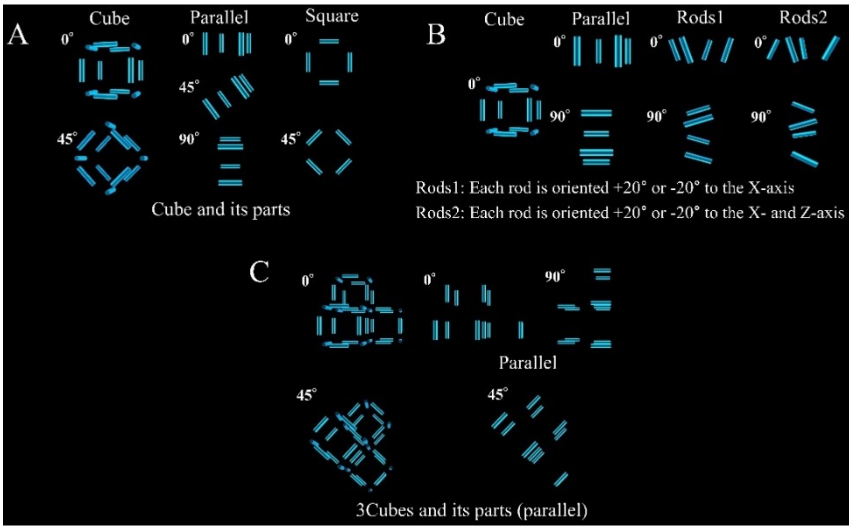


Fig. 1. Targets used in the experiments. The cube (A and B) or the three cubes (C) were the basic objects, and their parts were generated by removing several cylinders from the basic object. The orientations of these parts were varied.

Procedure

The participants wore the HMD provided and stood in the middle of the virtual 3D cubic space shown in Figure 2. Both the target and the distractors were presented in this virtual environment. A total of 1900 distractors were presented in the virtual environment. The target (Fig. 1) was presented at random positions on the x- and z-axes during each trial. The y-axis coordinates of the target were fixed.

When the participants held the controller in their hands and were ready, the target and distractors appeared on the HMD. The participants were tasked with finding and pointing out the predefined target (Fig. 1) among the distractors by pressing the pad of the controller as soon as they found the target. The participant could move their head, body or hand freely while searching. When the pad was pressed, the color of the target turned red for the feedback to the participant (Fig. 3). Determining whether a participant had correctly found the target was based on the direction to which the controller was pointed within the virtual environment. The reaction time (RT) associated with finding the target was also recorded. For the target set shown in Figure 1A, each participant completed 28 trials, which were presented randomly. All the trials included the presentation of one target out of seven (cube: 0° and 45°, parallel: 0°, 45°, and 90°, and square: 0° and 45°). The procedure was similar for the other target sets (Figs. 1B and 1C).

For comparison purposes, we also conducted the experiments in a standard manner in which the participants were seated in front of a standard 2D display on which the visual stimuli were presented.

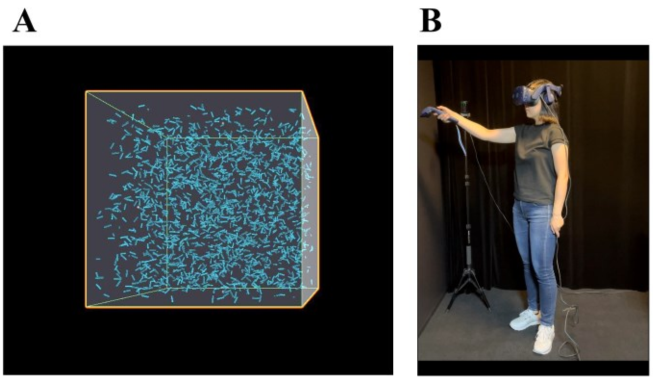


Fig. 2. A: Virtual environment with 1900 distractors. The initial position of the participants was at the center of this space. B: The participants could move their bodies freely.

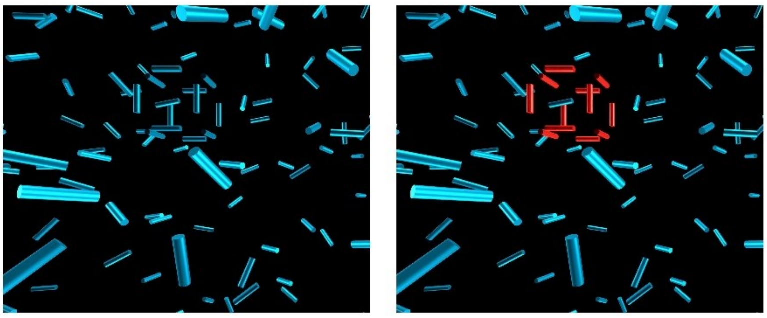


Fig. 3. Left image: Example of a participant’s view during a trial. The target in this case (a cube) is presented slightly above the center of the image. Right image: To ensure feedback to the participants, the targets turned red when the participants pressed the pad of the controller.

Results

Figure 4 shows the percent correct scores (Fig. 4A) and RT (Fig. 4B) as functions of the types of targets presented in Figure 1A. As shown in Figure 4A, the basic object (cube: 0°) was the easiest to search, but parallel: 0° (vertical components of the cube) was similarly detectable. The other targets were significantly less detectable than cube: 0° or parallel: 0°. The RTs (Fig. 4B) showed that the cube: 0° and the parallel: 0° were much faster to find compared to the rest of the targets. It should be noted that the levels of detectability between parallel: 0° and parallel 90° (horizontal components of the cube) were significantly different.

The Bonferroni correction method was used to confirm whether these considerations could be supported statistically. For percent correct scores (Fig. 4A), there were no significant differences among the cube: 0°, cube: 45°, or parallel: 0° (*ts* (15) ≦ 14.7, *n.s.*) and among the rest of the targets (*ts* (15) ≦ 2.9, *n.s.*). Significant differences between these two groups (*ts* (15) ≧ 4.1, *ps* < .02) were found. For the RTs (Fig. 4B), significant differences between parallel: 0° and parallel: 90° (*ts* (11) ≧ 4.1, *ps* < .03) were found.

The visual search performance was the highest for the basic object (cube: 0°), and that for the parallel parts was almost as high as that for the cube but only when they were vertically oriented (parallel: 0°). The difference in sensitivity levels between parallel: 0° (vertical) and parallel: 90° (horizontal) was remarkable, and this difference was not a result of the speed-accuracy trade-off because the RT was faster for parallel: 0° than that for parallel: 90°. This suggests that the detection of objects in noisy backgrounds depends on the vertical components or parts of such objects because their vertical parts would be processed to a higher degree compared to parts with other orientations.

Figure 5 shows the results for the targets shown in Figures 1B and 1C. Tendencies similar to those presented in Figure 4 were found. The percent correct for parallel: 0° (vertical components) was nearly the same as that for the main object (cube: 0° or three cubes: 0°). When the cylinders were slightly oriented (rods 1 or 2), the absolute performance decreased, but the superiority of vertical parts compared to that of horizontal parts was still observed. For the three-cube case, the sensitivity to horizontal components increased, but the increased sensitivity to vertical components was retained.

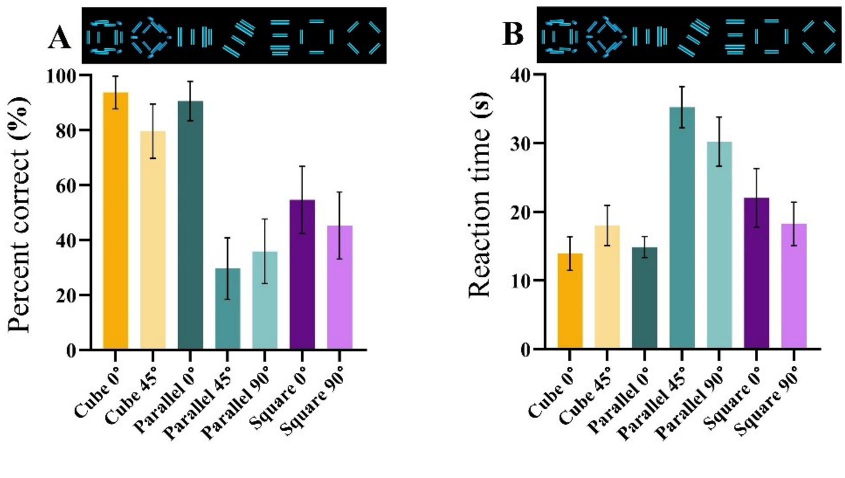


Fig. 4. Results for the targets shown in Fig. 1A and the upper insets of the graphs. Percent correct (A) or RT (B) as functions of the target types are presented. Error bars denote a 95% confidence interval.

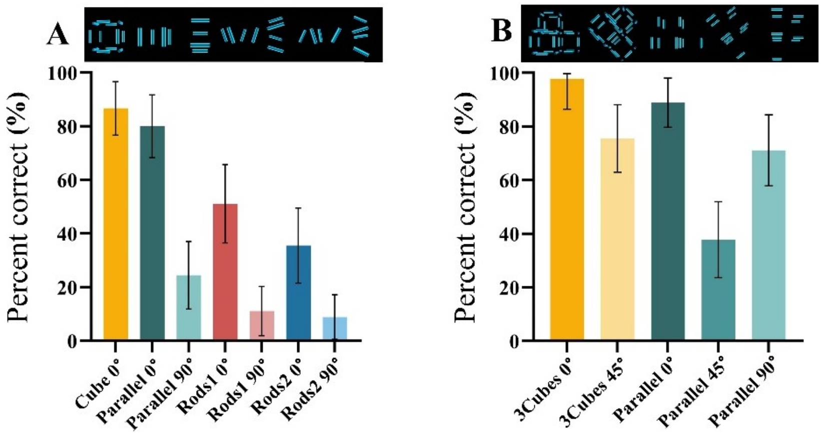


Figure 5. Percent correct (A) for Fig. 1B or percent correct (B) for Fig. 1C as functions of the target type are presented. Error bars denote a 95% confidence interval.

Discussion

We examined the visual search performance of a 3D object and its parts in a VR environment and established that there were high sensitivity levels toward vertical parts compared to those toward oblique or horizontal parts. As the visual search performance between the main object (cube) and its vertical parts was almost the same, we speculate that people depend on vertical parts during 3D object recognition.

From the other experiments, we established that the superiority of vertical targets in visual search performance compared to that of oblique or horizontal targets was more conspicuous in VR environments compared to that in standard 2D display environments. As the participants could move freely in the VR environment (Fig. 2), the vertical parts remained "vertical" in the spatiotopic reference frame. Therefore, high sensitivity levels toward the vertical parts of objects indicate that objects and their parts can be represented not only in retinotopic but also in spatiotopic reference frames, and our visual system can utilize the information obtained from the parts of objects in the spatiotopic reference frame for 3D object recognition. Since a slight deviation from the vertical position decreased the levels of sensitivity (Fig. 5A), the representation of vertical parts in the visual system is precise.

Why does vertical superiority exist? We speculate that it stems from the fact that people's daily lives are rich in vertical components. As shown in Figure 6A, throughout people's daily lives, standalone horizontal components in the height of their eyes are not conspicuous, except for artificial or industrial objects. However, people's daily lives are rich in vertical components, as shown in Figure 6A. This indicates that vertical components can be used to provide additional information regarding 3D objects compared to horizontal components. Meanwhile, the well-known oblique effect (high sensitivity to both vertical and horizontal components) (Heeley, et al., 1997) is reported to be observed in both retinotopic and spatiotopic reference frames depending on the conditions for observation (Mikellidou, et al., 2015). The relationship between the oblique effect and superior sensitivity to the vertical parts is yet to be understood.

The next question involves the way in which vertical components are extracted from input images. One possible mechanism involves low-level visual mechanisms, such as the association field model proposed by Field et al. (1993) (Fig. 6B), in which the association is postulated along the axis given by the neuron's orientation through the long-range excitatory and inhibitory connections. To examine this hypothesis, in our future studies, we plan to examine whether the subjective contours (Fig. 7A) can provide similar results, such as those presented in Figure 4. This is because the higher-order and not the lower-order visual mechanisms are responsible for inducing the impression of subjective contours (Kok, et al., 2016). Additionally, because we only used a simple basic figure (cube), in our future studies, we also plan to use complex basic figures, such as human-like shapes (Marr & Nishihara, 1978), as shown in Figure 7B.

Figure 6. A: Examples of vertical shapes in daily life. B: The association field model proposed by Field et al. (1993).

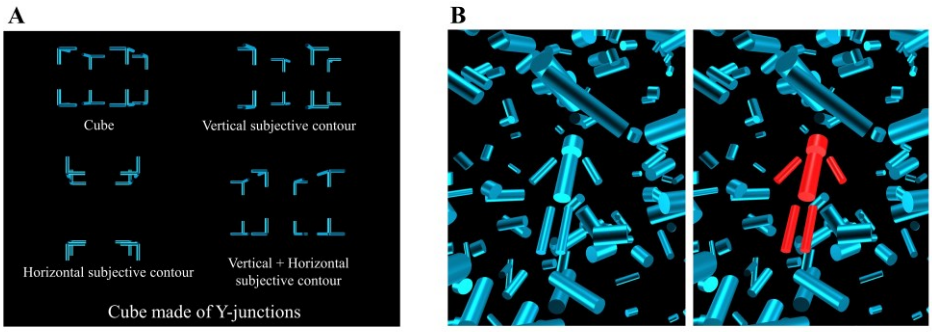


Figure 7. A: Targets defined using subjective contours. B: Target with a human-like shape.

Acknowledgments

This study was supported by JSPS KAKENHI: Grant Numbers JP21J12257 and 20H01788.

References

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115-147.

Field, D. J., Hayes, A. & Hess. R. F. (1993). Contour integration by the human visual system – Evidence for a local association field. *Vision Research, 33*, 173-193.

Heeley, D. W., Buchanan-Smith, H. M., Cromwell, J. A. & Wright, J. S. (1997). The oblique effect in orientation acuity. *Vision Research, 37*, 235-242.

Humphreys, G. W. & Donnelly, N. (2000). 3-D constraints on spatially parallel shape perception. *Perception & Psychophysics, 62*, 1060-1085.

Kok, P., Bains, L. J., van Mourik, T., Norris, D. G. & de Lange, F. (2016). Selective activation of the deep layers of the human primary visual cortex by top-down feedback. *Current Biology, 26*, 371-376.

Marr, D. & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London B., 200*, 269-294.

Mikellidou, K., Cicchini, G. M. & Thompson, P. (2015). The oblique effect is both allocentric and egocentric. *Journal of Vision, 15*, 1-10.

TESTING THE COMMON MECHANISM HYPOTHESIS RELATING PERCEPTUAL FILLING-IN TO MOTION INDUCED BLINDNESS DURING BINOCULAR RIVALRY

Rebecca A. White[1], Emily R. Komerska[1], Jonah K. Schneider[1], John E. Sparrow[2], and Wm Wren Stine[1]
[1]*Department of Psychology, University of New Hampshire, Durham, NH, USA*
[2]*Social Science Division, University of New Hampshire, Manchester, NH, USA*
<Rebecca.White@unh.edu, Bill.Stine@unh.edu>

Abstract

The visibility of a target region fluctuates in and out of conscious awareness when superimposed on a globally changing stimulus, known as motion induced blindness (MIB). When the target is filled in by a surrounding pattern rather than a uniform background, it is referred to as perceptual filling-in (PFI). These bistable, rivalrous phenomena are thought to be modulated by a common mechanism that is also shared with binocular rivalry (BR), referred to as the common mechanism hypothesis (Carter et al., 2003). In the present study, the common mechanism hypothesis is tested by inducing MIB/PFI during BR and measuring their temporal dynamics. We compare these dynamics to matched, non-rivalrous conditions, allowing us to explore potential interactions arising through this phenomenological cooccurrence. There were no differences across three measures between rivalrous and matched non-rivalrous conditions, suggesting that, though these phenomena seem to share many computational principles, they do not share a single oscillatory mechanism. Interestingly, there is a large significant difference between the MIB and PFI conditions, which replicates White et al. (2021). Further research is needed to parse the underpinnings of this finding.

MIB is a phenomenon characterized by one or more target regions in the periphery disappearing when superimposed on a globally changing pattern (e.g., motion; Bonnet et al., 2001). Understanding of the mechanisms driving MIB is limited, though it is thought to be related to a highly similar phenomenon called perceptual filling-in (PFI). Like MIB, PFI renders target stimuli stochastically invisible via an empty space in a pattern becoming filled in with that pattern (Ramachandran & Gregory, 1991). PFI and MIB are thought to share a common mechanism (Devyatko et al., 2016, Hsu et al., 2004, 2006), given their similarity in nature and that there are no differences in their dynamics when eccentricity, contrast, and size are controlled within the stimulus display. New and Scholl (2008, 2018) hypothesize that MIB and PFI are the same phenomenon under different perceptual conditions, functioning to enhance visual clarity by expunging the target stimulus as if it were an artificial scotoma or an artificial optic disk, which are normally filtered from awareness by the visual system. Essentially, the scotoma theory posits that the visual system is exploring two equally likely interpretations of the sensory data: Does the target exist in the distal world, or it is an artifact of the eye?

Images that cannot be binocularly fused stochastically oscillate in conscious awareness, referred to as binocular rivalry (BR; von Helmholtz, 1929). BR, MIB, and PFI all constitute bistability – the stimuli remain constant, but the observer's perceptual state switches between two sensory interpretations (i.e., in BR, the two images, and in MIB/PFI, either *"target visible"* or *"target invisible"*). MIB can be thought of as global versus local rivalry, while BR can be thought of as global versus global rivalry. In 2003, Carter and Pettigrew proposed that BR and MIB share a common oscillatory mechanism expressed within the pontine brainstem, given the remarkable temporal similarities within individuals and that they share identical differences when stimulating the left versus right hemisphere with transcranial magnetic stimulation (TMS; Funk & Pettigrew, 2003). Further, reversal rate of both phenomena is slowed by administration of psychedelic drugs (Carter et al., 2005, 2007) and in certain clinical conditions, such as psychosis, autism, or obsessive-compulsive disorder (Robertson et al., 2013; Tschacher et al., 2006; Vierck et al., 2013; Ye et al., 2013).

Via dichoptically presenting an MIB target rivalrous in color, Jaworska and Lages (2014) found that certain MIB manipulations (e.g., target size) had no effect on BR, and vice versa, which constitutes counterevidence for the common mechanism hypothesis – MIB and BR must involve differential processing while competing for visual awareness. Given that the targets were only rivalrous in color and not in shape/form, which possess different mechanisms of processing (e.g., Holmes et al., 2006), White et al. (2021) conducted a similar experiment wherein line stimuli of opposing orientation induced BR and MIB simultaneously. When orientations were matched in both eyes (i.e., when the stimuli were nonrivalrous), there was no difference in the temporal dynamics compared to when the orientations did induce BR. This further supports Jaworska and Lages, suggesting that the mechanisms responsible for MIB and BR are independent systems. As well, White et al. induced BR and PFI simultaneously via line orientations within grid stimuli. Curiously, there was a small effect of BR within PFI displays. There was also a large significant difference between the temporal dynamics of MIB and PFI, which is counterintuitive given that the target areas were matched in terms of contrast, size, and eccentricity, and New and Scholl (2008; 2018) go so far as to suggest that they are the same phenomenon. Both of these surprising findings were interpreted as most likely being attributable to an artifact of exploratory stimuli – the area being filled in was not surrounded by what it must be filled in by, unlike MIB, and thus they were not truly matched (see White et al., 2021 for stimulus displays). Further, participants reported that it was difficult to tell when filling in was occurring, which likely explains the small significant difference between rivalrous and nonrivalrous conditions in PFI. The present study replicates White et al., 2021, using cosine wave grating stimuli rather than lines/grids.

Methods

Participants

Four volunteers with normal or corrected to normal vision were recruited. Participants signed informed consent forms and received debriefing consistent with University of New Hampshire Institutional Review Board policy. There was no compensation.

Apparatus & Stimuli

See White et al., 2021 for monitor specs, information on the Maxwellian View optical system, and descriptions of the motion mask and nonius line components of the stimulus display. Stimuli were created using the PsychToolbox extensions (Brainard, 1997; Pelli, 1997) in *MatLab* (v. 2019b, MathWorks). PFI trials included a cosine wave grating subtending .62º with the center placed at .85º eccentricity relative to the nonius lines. The center of the cosine wave grating was covered by a circle of 7.4 arcmin filled with the same shade of grey as the background. Spatial frequency of the grating was 19.36 c/deg (i.e., 12 cycles) and Michelson contrast was .94 Td of retinal illuminance. MIB stimuli included a circle of identical size and eccentricity filled with the cosine wave grating (see Figure 1). ~2 cycles can be seen within the MIB target.

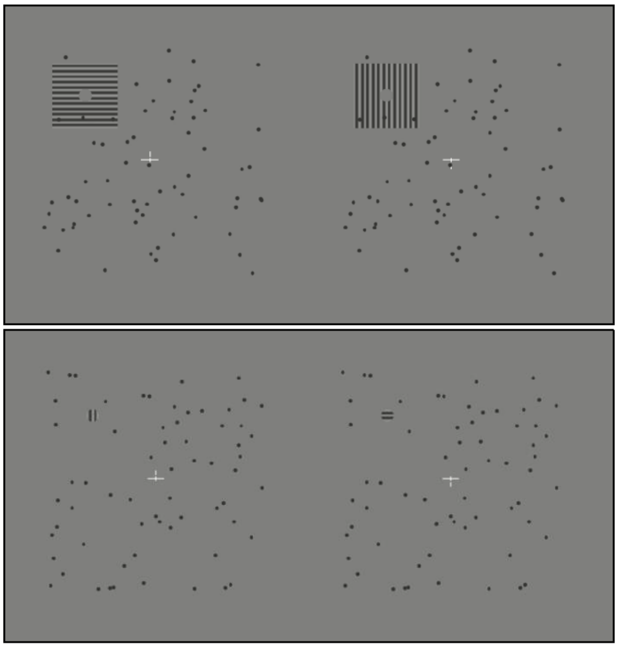


Figure 1. Top: Rivalrous PFI display. Bottom: Rivalrous MIB display.

Procedure

The grating's orientation (horizontal or vertical) varied across trials, either matching in both eyes, or opposite in each eye. Each stimulus set was randomly presented eight times. The experiment took place in a darkened room and was preceded by a 120 second adaptation period. Effects were indicated by participants via depressing the spacebar on a wireless keyboard sampled at 60 Hz. See White et al., 2021 for a more detailed procedure description.

Results

Effects were analyzed using a randomized block factorial subject by stimulus (6) and trial (8) repeated measures analysis of variance (randomized block factorial design, RBF-68; Kirk, 2013, Ch. 10). Exploratory data analyses revealed minor deviations from normality, which were corrected for using an arcsine transform. All critical F and p values were adjusted using the Chi Muller statistic and the treatment by block interactions were explored, given that the assumptions of sphericity and additivity were violated when tested using the locally best invariant test and Tukey's test of nonadditivity, respectively. To control familywise Type I error rate of .05, a critical alpha level of .016 was used across the three dependent variables included in the model: cumulative duration of the effect, duration to the onset of the effect, and perceptual reversals (i.e., the number of target disappearances and reappearances). All pairwise comparisons were tested *a posteriori* using the SAS Simulate procedure (*SAS* v9.4), which uses the control variate adjustment method of Hsu and Nelson (1998) to control for the familywise Type I error rate of .05.

There was a significant effect of stimulus across all measures: cumulative duration, *CM Adj.* $F(0.5, 3) = 36.19$, $p = .003$, partial $\omega^2 = .49$; duration to onset, *CM Adj.* $F(0.2, 1) = 14.57$, $p = .02$ partial $\omega^2 = .26$; and reversals, *CM Adj.* $F(0.7, 4) = 19.28$, $p = .001$, partial $\omega^2 = .49$. Pairwise comparisons revealed significant differences between MIB and PFI effects across all measures (all p's < .001) such that there was a stronger effect of MIB compared to PFI, which replicates the findings of White et al., 2021. There were no significant differences between rivalrous and matched stimuli and no effect of trial.

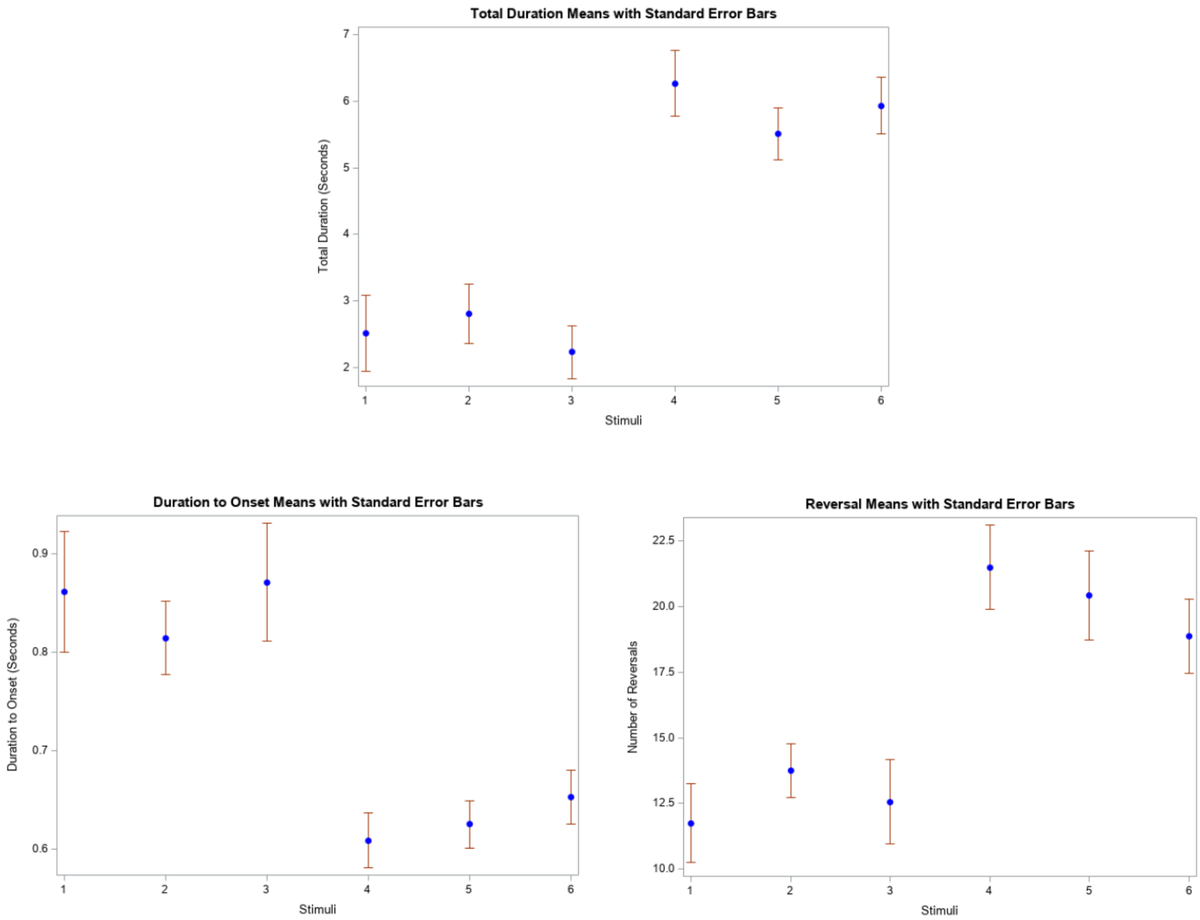


Figure 2. Means with standard error bars. Stimuli 1-3 represent PFI; Stimuli 4-6 represent MIB. Stimuli 1 and 4 represent the rivalrous conditions, others represent matched conditions.

Discussion

The present study investigated MIB and PFI dynamics during BR to further probe the common mechanism hypothesis test whether the findings from White et al. 2021 would replicate with improved stimulus designs. As expected, the significant difference between the rivalrous and non-rivalrous conditions in PFI was no longer present, which further supports Jaworska and Lages' (2014) counterevidence to the common mechanism hypothesis. In line with this, participants vocalized greater ease in discriminating when PFI was occurring compared to the previous experiment, and they were instructed to only report the effect when the target was clearly and fully filled in. With respect to the common mechanism hypothesis, MIB, PFI, and BR likely share many computational principles, but not a single common oscillator.

A large significant difference remains between MIB and PFI conditions. It is unclear whether this finding reflects a genuine difference between the phenomena, or whether it is an artifact of the stimulus design. Given that the target regions were matched in contrast, size, and eccentricity, this finding remains an intriguing anomaly which should be explored in future studies (Hsu, Yeh, & Kramer, 2006).

References

Bonneh, Y. S., Cooperman, A., & Sagi, D. (2001). Motion induced blindness in normal observers. *Nature*, *411*(6839), 798–801. https://doi.org/10.1038/35081073

Brainard, D. H. (1997). The psychophysics toolbox. Spatial Vision, 10, 433–436.

Carney, T., Shadlen, M., & Switkes, E. (1987). Parallel processing of motion and colour information. *Nature, 328*(6131), 647-649. https://doi:10.1038/328647a0

Carter, O. L., & Pettigrew, J. D. (2003). A common oscillator for perceptual rivalries? *Perception, 32*(3), 295-305. https://doi:10.1068/p3472

Carter, O. L., Pettigrew, J. D., Hasler, F., Wallis, G. M., Liu, G. B., Hell, D., & Vollenweider, F. X. (2005). Modulating the rate and rhythmicity of perceptual rivalry alternations with the mixed 5-HT2A and 5-HT1A agonist Psilocybin. *Neuropsychopharmacology, 30*, 1154-1162. https://doi:10.1038/sj.npp.1300621

Carter, O. L., Hasler, F., Pettigrew, J. D., Wallis, G. M., Liu, G. B., & Vollenweider, F. X. (2007). Psilocybin links binocular rivalry switch rate to attention and subjective arousal levels in humans. *Psychopharmacology,* (195), 415-424. https://doi:10.1007/s00213-007-0930-9

Devyatko, D., Appelbaum, L. G., & Mitroff, S. R. (2016). A common mechanism for perceptual reversals in motion induced blindness, the Troxler effect, and perceptual filling-in. *Perception, 46*(1), 50-77. https://doi:10.1177/0301006616672577

Funk, A. P., & Pettigrew, J. D. (2003). Does interhemispheric competition mediate motion-induced blindness? A transcranial magnetic stimulation study. *Perception, 32*(11), 1328-1338. https://doi:10.1068/p5088

Holmes, D. J., Hancock, S., & Andrews, T. J. (2006). Independent binocular integration for form and colour. *Vision Research, 46*(5), 665-677. https://doi:10.1016/j.visres.2005.05.023

Hsu, J. C., & Nelson, B. (1998). Multiple comparisons in the General Linear Model. *Journal of Computational and Graphical Statistics, 7*(1), 23-41. https://doi:10.1080/10618600.1998.10474759

Hsu, L.C., Yeh, S.-L., & Kramer, P. (2004). Linking motion induced blindness to perceptual filling-in. *Vision Research*, 44, 2857 – 2866. https://doi:10.1016/j.visres.2003.10.029

Hsu, L.C., Yeh, S.-L., & Kramer, P. (2006). A common mechanism for perceptual filling-in and motion-induced blindness. *Vision Research*, 46, 1973– 1981. https://doi:10.1016/j.visres.2005.11.004

Jaworska, K., & Lages, M. (2014). Fluctuations of visual awareness: Combining motion-induced blindness with binocular rivalry. *Journal of Vision, 14*(11), 11. https://doi:10.1167/14.11.11

Kirk, R. W. (2013). Experimental design: Procedures for the behavioral sciences (4th ed.). Washington, DC: Sage.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.

New, J. J., & Scholl, B. J. (2008). "Perceptual Scotomas." *Psychological Science*, *19*(7), 653–659. https://doi.org/10.1111/j.1467-9280.2008.02139.x

New, J. J., & Scholl, B. J. (2018). Motion-induced blindness for dynamic targets: Further explorations of the perceptual scotoma hypothesis. *Journal of Vision*, *18*(9), 24. https://doi.org/10.1167/18.9.24

Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature, 350*(6320), 699-702. https://doi:10.1038/350699a0

Robertson, C. E., Kravitz, D. J., Freyberg, J., Baron-Cohen, S., & Baker, C. I. (2013). Slower rate of binocular rivalry in autism. *The Journal of Neuroscience, 33*(43), 16983-16991. https://doi:10.1523/jneurosci.0448-13.2013

White, R. A., Sparrow, J. E., Dickinson, K., Tremblay, D., LaBarre, J. A., Kitt, A. J., & Stine, W. W. (2021). The dynamics of perceptual filling-in and motion induced blindness during binocular rivalry: Exploring the common mechanism hypothesis. In J. R. Schoenherr, T. Hubbard, & K. Ueda (Eds*.), Proceedings of the 37th annual meeting of the international society for psychophysics: Fechner day 2021* (pp. 68-71). International Society for Psychophysics.

Ye, X., Zhu, R., Zhou, X., He, S., & Wang, K. (2019). Slower and less variable binocular rivalry rates in patients with bipolar disorder, OCD, major depression, and schizophrenia. *Frontiers in Neuroscience, 13*. https://doi:10.3389/fnins.2019.00514

INFORMATION PROCESSING ABNORMALITIES AMONG PATIENTS WITH SCHIZOPHRENIA IN A DECISION-MAKING TASK

Alexandra Wolf[1, 2], Shunsuke Tamura[2], Takako Mitsudo[2], Kazuo Ueda[3], and Yoji Hirano[2]

[1]*Cognitive Behavioral Assistive Technology Team, RIKEN Center for Advanced Intelligence Project, Tokyo 103-0027, Japan*

[2]*Department of Neuropsychiatry, Graduate School of Medical Sciences, Kyushu University, Fukuoka 812-8582, Japan*

[3]*Department of Acoustic Design/Research Center for Applied Perceptual Science/Research and Development Center for Five-Sense Devices, Kyushu University, Fukuoka 815-8540, Japan*

<alexandra.wolf@riken.jp>

Abstract

Schizophrenia (SZ) is a severely debilitating mental disorder characterized by disturbances in thought, perception, and behavior. This disorder ranks among the top 10 disabling conditions worldwide among young adults who experience uncontrolled anxiety and progressive impairment of social skills. As a result, patients with SZ may experience the inability to lead an independent and fulfilling life (compared to healthy controls, HC). An ample of clinical investigations have reported that when compared to eye movements of cognitively intact participants, the gaze behavior of patients with first-episode and chronic SZ is abnormal (for a systematic review, kindly refer to Wolf et al., 2021). For example, patients with SZ tend to shift their fixation less frequently, i.e., gazing at only one area of the presented stimulus. In comparison, HC view pictures more widely. Although the eye-tracking methodology has existed in the clinical domain for a long time, there is an apparent deficiency of investigations that combine this technology with decision-making paradigms. The presented framework is a preliminary study among patients with SZ on their evaluative responses toward naturalistic food images (kindly see Wolf et al., 2018 & 2019). To understand the real-life interplay between cognition and eye behavior, the authors introduce ecologically valid situations, i.e., "Liking" and "Shopping." In the "Liking" task, participants were asked to rate each single-displayed food image from 1 ("not like at all") to 3 ("like very much"). In the "Shopping" task, participants were instructed to put a maximum of 7 food images into their shopping baskets. As in the "Liking" task, the number of potential response options was three. Hence, in the "Shopping" task, the possible responses were: "leave it", "postpone judgment", and "put in". The relationship between rating and viewing time differed between patients with SZ (N=24) and HC (N=30). For example, in the "Shopping" task, a significant inverted U-shape trend has been observed among HC (i.e., "postponed" food images were associated with the longest viewing time). Such a relationship could not be observed among patients with SZ, who faced indecisiveness while being requested to make a more subjective decision. Experimental tasks that mirror real-life situations may be useful in future comparisons across various clinical populations.

THE EFFECTS OF THE SELECTION AND HYPERACUSIS ON THE PERCEPTION OF BACKGROUND MUSIC

Teruo Yamasaki
Department of Psychology, Osaka Shoin Women's University, Higashi-osaka, Osaka, Japan
<yamasaki.teruo@osaka-shoin.ac.jp>

Abstract

Background music is widely used both in public and private spaces. It can have a variety of effects, such as emotionally, physically, cognitively, and interpersonally beneficial effects. Many factors could contribute to such effects, among them, the perceived characteristics, particularly pleasantness or liking of background music is a crucial factor. The perceived characteristics of background music depend not only on the properties of the music itself but also the environment in which it is played and the characteristics of the listeners. Among the environmental factors, the selection control of background music, namely who chooses the musical pieces played, may influence its perceived characteristics. The sound sensitivity of listeners, namely hyperacusis, may also influence the characteristics of background music. In this study, 320 respondents (160 males and 160 females, mean age =39.7, and SD =10.9) answered a questionnaire which included Khalfa's Hyperacusis questionnaire and questions about the characteristics of background music, particularly pleasantness, in public spaces, private spaces, and the workplace; the degrees of selection control are different in each of these spaces. Analysis of the data revealed that (1) pleasantness of background music differed between public and private spaces; this is related to the difference of selection control in both spaces. (2) Sound sensitivity was negatively correlated to the pleasantness of background music in public spaces but did not show such a strong correlation to the pleasantness of background music in private spaces. (3) Pleasantness of background music in workspaces was positively correlated to sound sensitivity, although liking of background music in public and private spaces was negatively correlated to sound sensitivity. These results suggest that selection control and sound sensitivity influence the perception of background music.