Contents lists available at ScienceDirect

Cognition



journal homepage: www.elsevier.com/locate/COGNIT

Thresholds for color discrimination in English and Korean speakers

Debi Roberson^{a,*}, J. Richard Hanley^a, Hyensou Pak^b

^a University of Essex, Department of Psychology, Wivenhoe Park, Colchester CO4 3SQ, UK ^b Yeungnam University, South Korea

ARTICLE INFO

Article history: Received 17 September 2008 Revised 9 April 2009 Accepted 10 June 2009

Keywords: Color categories Cultural relativity Discrimination thresholds Just-noticeable-difference Perception

ABSTRACT

Categorical perception (CP) is said to occur when a continuum of equally spaced physical changes is perceived as unequally spaced as a function of category membership (Harnad, S. (Ed.) (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. Cambridge: Cambridge University Press). A common suggestion is that CP for color arises because perception is qualitatively distorted when we learn to categorize a dimension. Contrary to this view, we here report that English speakers show no evidence of lowered discrimination thresholds at the boundaries between blue and green categories even though CP is found at these boundaries in a supra-threshold task. Furthermore, there is no evidence of different discrimination thresholds between individuals from two language groups (English and Korean) who use different color terminology in the blue–green region and have different supra-threshold boundaries. Our participants' just noticeable difference (JND) thresholds suggest that they retain a smooth continuum of perceptual space that is not warped by stretching at category boundaries or by within-category compression. At least for the domain of color, categorical perception appears to be a categorical, but not a perceptual phenomenon.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Categorical perception (CP) for color is manifest as more accurate or more rapid discrimination of pairs of colors that belong to different color categories compared to equally separated members of the same category (when equated in a smooth perceptual metric such as C.I.E. L^{*}u^{*}v^{*} space. For example, participants can detect a difference between a blue target and green distractors more quickly than they can detect a difference between targets and distractors that belong to the same color category (e.g. different shades of blue). They also show faster peak ERP latencies for between- compared to within-category color differences (Fonteneau & Davidoff, 2007), which suggests that categorization takes place within 100 ms and before visual analysis is complete. One explanation of CP is

E-mail address: robedd@essex.ac.uk (D. Roberson).

0010-0277/\$ - see front matter @ 2009 Elsevier B.V. All rights reserved. doi:10.1016/j.cognition.2009.06.008

that our perception is warped so that an otherwise smooth continuum of change becomes 'stretched' at category boundaries and 'compressed' in category centers (Harnad, 1987; Notman, Sowden, & Ozgen, 2005; Özgen & Davies, 2002; Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009). In principle, such perceptual inequality might arise from the innate structure of human color vision. However adult speakers of languages that employ fewer (e.g. Berinmo, Himba) or more (e.g. Russian, Korean, Greek) basic color categories than English show CP in different areas of color space from speakers of English (Athanasopoulos, 2009; Davidoff, Davies, & Roberson, 1999; Roberson, Davidoff, Davies, & Shapiro, 2005; Roberson, Pak, & Hanley, 2008; Winawer et al., 2007). It would therefore follow that any perceptual inequality that exists for these individuals must arise because category learning leads to changes in perceptual systems. Such changes would involve relatively long-term tuning at a receptor level rather than short-term adaptation, attentional shifts, or changes in strategic focus (Goldstone, 1998). If this is true, then it should be possible



Brief article

^{*} Corresponding author. Tel.: +44 01206 874175/37; fax: +44 1206 873590.

to discriminate just noticeable differences (JNDs) between shades of color more readily at boundaries between color categories than in category centers.

Nevertheless, some evidence rests uneasily alongside claims that CP for color is caused by perceptual warping. CP for color is not observed when naming is prevented by a secondary verbal task (Gilbert, Regier, Kay, & Ivry, 2006; Roberson & Davidoff, 2000; Winawer et al., 2007). Furthermore, when targets are presented either to the left or right of a central fixation point, color CP is observed in the right visual field (RVF) but not the left (Drivonikou et al., 2007; Roberson et al., 2008). Such hemispheric asymmetry has been shown for English speakers at the boundary between blue and green (Gilbert et al., 2006), and for Korean speakers for a boundary that is marked in Korean, but not in English (Roberson et al., 2008). The asymmetry seems to occur because stimuli presented in the RVF have preferential access to language processing areas in the left hemisphere. Further evidence comes from the finding of the asymmetry in a split-brain patient, for whom information about stimuli presented in the left visual field could not reach the left hemisphere (Gilbert et al., 2006). Moreover, recent evidence shows differential activation of brain areas involved in language processing during early perceptual processing of color (Tan et al., 2008).

These findings raise the possibility that CP for color is mediated by higher-level cognitive processes rather than by perceptual warping, and only occurs when a linguistic code is accessed. According to this view, CP occurs because colored stimuli from different color categories are represented as separate terms (e.g. blue vs. green) in a verbal code. Consequently, they can be distinguished rapidly because both verbal and perceptual codes provide converging evidence that they are different. Two different shades of the same color category will be distinguished more slowly because they will activate the same verbal label, which will conflict with the perceptual information that they are different (Roberson & Hanley, 2007; Roberson et al., 2008). According to this view, while CP should be observed on supra-threshold tasks, it will not be observed on a task that involves discrimination of INDs because its performance is unlikely to be affected by activation of a verbal code.

As a direct test of the perceptual warping hypothesis, the present study compares JND discrimination thresholds (the smallest difference in shade that can be reliably discriminated) in speakers of English and Korean. We used a range of colors that English speakers divide into two categories (green and blue) but Korean speakers divide into three basic categories: *chorok, cheongnok* and *parang* (Kim, Pak, & Lee, 2001; Lee et al., 2003).¹ CP for this range of stimuli has previously been demonstrated in English speakers for a wide range of supra-threshold perceptual and memory judgments (Bornstein & Korda, 1984; Pilling, Wiggett, Özgen, & Davies, 2003; Roberson & Davidoff, 2000; Wiggett & Davies, 2008). We chose to use the ZEST algorithm (King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994) because it avoids the problems associated with traditional adaptive staircase methods that systematically increase or decrease the difference being measured. In those methods participants can anticipate the mechanics of the staircase. This is not possible with the ZEST algorithm, which is not systematic in its increments or reductions of color separations. We also compare the performance of Korean and English speakers on a supra-threshold same-different judgment task (Experiment 2). If learning the appropriate set of categories warps perceptual space, then discrimination thresholds should be lower at the boundary between an individual's named categories than in the category centers. These lowered thresholds should be found at those boundaries for which CP is found in supra-threshold tasks. There should also be group differences between English and Korean speakers relating to the number of categories used and the location of category boundaries.

2. Experiment 1 – discrimination thresholds

To test discrimination thresholds, eight monolingual native English speakers (tested in Colchester, UK) and eight native Korean speakers (tested in Seoul, Korea) carried out an edge-detection task for patches of color presented on a computer screen in a dark room (luminance less than 0.001 cd/m²). All participants were tested in their native language and had no knowledge of color terms in the other language. All had normal color vision as assessed by the Ishihara (1992) and City Color Vision (Fletcher, 1980) tests. Two adjacent rectangular patches of color appeared in the centre of the screen on a neutral grey background, forming a rectangular patch of which 2/3 was one shade and the other 1/3 a slightly different shade. The 'edge' between two adjacent patches of color was blurred with a 6 pixel wide Gaussian blur and could be reliably detected only if the two shades could be discriminated. Participants were asked to estimate whether a line (edge) appeared towards the left or right side of the patch.

Using the ZEST algorithm (King-Smith et al., 1994) participants' IND thresholds were estimated in CIE (1976, $L^*u^*v^*$) ΔE units for each of 15 points (linearly spaced in the perceptual metric) ranging from the centre of the green/chorok category, across the boundary between green and blue for English speakers, or across the boundaries between *chorok/cheongnok/parang* to the centre of the blue/ parang category for Korean speakers. These points were derived from Munsell colors, for which previous experiments have established the boundary for each of the above categories and maintained lightness (Munsell Value = 5.47) and saturation (Munsell Chroma = 7.03) constant, so they varied only in hue. They were converted to L^{*}u^{*}v^{*} (Appendix I) and implemented using the CRS VISage 11/3 protocol on a Mitsubishi, Diamond Pro 2070, 22" CRT monitor (set to gamma - 2.2, color temperature - 6500 K, luminance -100 cd/m²). Color patches measured 7.5 cm \times 15.5 cm and were viewed from approx. 60 cm at eye-level. Luminance values varied minimally across the range (less than 10% ΔE). The resolution and reproduction frequency of the monitor were 1024×768 pixels and 85 Hz,

¹ These categories are basic according to the criteria of Kay, Berlin, & Merrifield (1991), being monolexemic, not subsumed under the meaning of other terms, not restricted to a narrow class of objects, and understood by all observers.

respectively and stimulus measurements were verified with a CRS ColorCal colorimeter. Using ZEST procedures, there were three interleaved runs in each cycle, each for one of the 15 target locations, with initial offset of ±.055 ΔE . Across trials the separation between target and distractor stimuli varied between .09 and .02 ΔE . Thresholds were bi-directionally measured around each point. Stimuli remained on screen for 1000 ms or until a response was made. The complete set of 5 cycles was presented twice in random order and the combination of points for which thresholds were measured in each cycle was randomized across participants so that thresholds were measured for each point in random order. For each stimulus the position of the target (left or right) and the fraction of the total patch it occupied (2/3 or 1/3) were randomly varied across trials and across comparison stimuli. Total testing time averaged 1 h. After testing, each participant named each of the 15 stimuli for which thresholds had been measured. To control for any effects of the lengthy threshold measurement task on naming, participants returned to



Fig. 1. Mean naming (a) for English speakers and (b) for Korean speakers across the range of stimuli tested from the centre of the green category to the centre of the blue category and (c) JND discrimination thresholds (in ΔE units) for English and Korean speakers across the same range of stimuli (with standard error bars).

the laboratory after a two-week interval and named the stimuli again. Comparisons were first made for each participant between thresholds for shades that were central to each of their own named categories and for shades that were on the boundary between their named categories. Subsequently, group comparisons were made between languages for thresholds at the English boundary and at the two Korean boundaries. Fig. 1 shows mean naming boundaries averaged across the two naming sessions for English speakers (a), Korean speakers (b) and mean threshold measurements across the range of stimuli for English and Korean speakers (c).

A 2: language (Korean vs. English) \times 2: stimulus type (Within-category vs. Boundary) ANOVA with repeated measures over the second factor revealed no significant difference in discrimination sensitivity for boundary relative to within-category stimuli for either population $[F_{(1,14)} < 1]$, no significant difference in sensitivity between speakers of the two languages $[F_{(1,14)} = 1.68, p > .1]$ and no interaction $[F_{(1,14)} = 1.69, p > .1]$. Due to the properties of the monitor there were minor perturbations in the differences of thresholds tested so that not all ΔE differences between a target and its nearest neighbors were exactly identical (maximum perturbation = 0.007 ΔE). These differences were extremely small and could not have affected the result, since the largest differences were measured for the boundary stimuli 7.5BG and 8.5BG. Had this marginally greater distance resulted in easier discrimination, we should have observed a reduced threshold at this boundary, but we did not. However, an additional control analysis compared thresholds at 3.7BG and 6.25BG (withincategory green and blue, respectively) with thresholds at 7.5BG and 8.75BG (cross-category) for which differences from closest comparison stimuli were identical (0.02 ΔE). The ANOVA revealed no significant effect of language $[F_{(1,14)} = 2.03, MSE = 1.91, p > .1]$, no significant effect of pair type $[F_{(1,14)} < 1]$ and no significant interaction $[F_{(1,14)} < 1]$. Finally, we compared performance for English and Korean speakers at the English boundary (7.5BG), and at the 2 Korean boundaries (10BG and 3.75BG). There was no significant difference between the two populations for any of the measured points (Bonferroni $t_{(14)} = .331, 1.77$ and 1.57, respectively, all p > .05) although there was a slight trend for Korean speakers to have lower thresholds than English speakers for the two Korean boundaries. No participants showed a significantly reduced discrimination threshold at the category boundary, in an edge-detection task that should not involve color naming.

The results clearly show no systematic differences in discriminative sensitivity at the boundaries between English color categories or between speakers of Korean and English at category boundaries that exist in Korean but not English. To confirm that both English and Korean speakers show CP at these boundaries in supra-threshold tasks, Experiment 2 compared same-different judgments for the 15 stimuli for which thresholds were measured in Experiment 1.

3. Experiment 2

From the 15 target stimuli in Experiment 1, within- or across-category pairs of equally separated colors were created and matched with an equal number of pairs of identical stimuli. Twenty native English speakers (7 males, 13 females, mean age = 22.3) from the University of Essex and 20 native Korean speakers (9 males, 11 females, mean age = 25.7) from higher-education institutions in Seoul, all with normal color vision made same-different judgments for these pairs of stimuli. None had taken part in Experiment 1. Equally spaced pairs of stimuli were constructed so that pairs 5BG/7.5BG and 6.25BG/8.25BG were withincategory for Korean speakers but cross-category for English speakers. Pairs 10BG/2.5B and 8.25BG/1.25B were crosscategory for Korean speakers but within-category for English speakers. Pair 3.75BG/6.25BG was cross-category for speakers of both languages and pair 5B/7.5B was withincategory for speakers of both languages. Each of the stimuli used in 'different' pairs also appeared paired with an identical patch as a 'same' pair. The 6 'different' pairs appeared twice, with each stimulus appearing once on the left and once on the right. Each block thus had 12 'same' and 12 'different' pairs. There were 4 blocks of trials with stimulus



Fig. 2. Mean proportion of correct discriminations (with standard error bars) of pairs of different colors that were either within-category in both languages; cross-category in English, within-category in English or cross-category in both languages.

pairs randomized within each block, making a total of 96 trials. Participants completed 8 initial practice trials with colors not included in the subsequent experiment. Color patches measured $5 \text{ cm} \times 5 \text{ cm}$, separated by 1.5 cm and were viewed from approx. 60 cm at eye-level in the centre of a neutral grey screen (properties of the screen given for Experiment 1). Stimuli were generated using the CRS VISage system and the same monitor as Experiment 1. Following a fixation cross for 700 ms, pairs of stimuli were presented for 350 ms, followed by a 750 ms ISI. Participants indicated whether the two patches were 'identical' or 'different' shades by key press.

The proportion of correct 'different' judgments was compared across languages in a 2 (Language: English vs. Korean) \times 4 (Pair type: within-category in both languages vs. within-category in English vs. cross-category in English vs. cross-category in both languages) ANOVA with repeated measures over the second factor. There was no significant effect of language $[F_{(1,38)} = 1.07, MSE = .01, p > .1]$ and no significant effect of pair type $[F_{(3,114)} = 1.63,$ MSE = .01, p > .1] but a significant interaction $[F_{(3,114)} =$ 7.75, MSE = .01, p < .01]. Newman–Keuls pairwise comparisons of the interaction showed a significant effect of Language for pairs that were within-category in English (cross-category in Korean) and cross-category in English (within-category in Korean) (both p < .01) but no difference for the other two pair types. English speakers showed significantly better discrimination for both types of crosscategory pairs than for either type of within-category pair (both p < .05). Korean speakers, however, showed significantly better discrimination of pairs that were within-category in English than pairs that were cross-category in English (p < .05). Fig. 2 illustrates these results.

For each language, pairs of colors that crossed the category boundary were discriminated more accurately than pairs of colors from within the same category, even though the boundaries differed across languages.

These findings support the findings of Roberson et al. (2008) for the Korean category boundary between *yeondu* (yellow–green) and *chorok* (green), which lies fully within the green category for English speakers. Roberson et al. (2008) suggested that the differential CP effect observed in speakers of different languages arise as a result of conflict for within-category decisions between access to the category label (which indicates that two stimuli are the same) and access to the visual information that they are different. Resolving these two conflicting sources of information leads to more error–prone performance than for cross-category pairs, where both the category label and the visual information indicate that the two stimuli are different.

4. General discussion

In Experiment 2 all participants showed CP for the boundaries marked in their own language, with more accurate discrimination of cross- than within-category decisions. This could not have been the result of any perceptual inequality between pairs of stimuli, because the pattern of performance differed across language groups. In Experiment 1, by contrast, neither English nor Korean participants showed lowered perceptual thresholds at the boundary between their named categories. The data strongly suggest that learning to categorize colors does not 'warp' color perception and that CP for color does not result from such perceptual warping.

In supra-threshold tasks, such as simultaneous samedifferent judgments, within- and cross-category discriminations might be equivalent because performance for both types of stimuli is at ceiling. This cannot be the case in Experiment 1. All participants do show variability in their discrimination thresholds across the range of stimuli tested. However this variation is not systematic and does not yield lower thresholds at individual named category boundaries. We conclude that, at least at the very finely graded level of Just Noticeable Differences, color space is perceptually uniform.

In an edge-detection task such as the present one, where the difference between the two shades is at the level of a IND, verbal labels would not aid discrimination. Although the whole series progresses from green to blue (for English speakers), the difference between any two stimuli is so small that they would not reliably be called by different names (see Roberson, Davidoff, & Braisby (1999) for a discussion of this issue). Unlike the same-different task used in Experiment 2 in which color CP was demonstrated, the threshold task does not require overt comparison of differences or similarities between colors. It therefore appears that the category advantage termed CP only occurs in tasks that involve activation of verbal categories or linguistic codes. When naming is of no benefit, as in the present task, or where it is prevented by verbal interference or by LVF presentation, there is evidence of a smooth continuum of change across the color spectrum.

In other domains where novices are trained to discriminate previously indistinguishable visual stimuli, it is possible that the development of expertise entails tuning at a more perceptual level. For example, improvement in some visual discriminations, such as the motion of a random dot field is associated with significant changes in the response characteristics of individual cells in parietal (Zohary, Celbrini, Britten, & Newsome, 1994) or primary visual cortex (Fahle & Morgan, 1996). Also, participants trained on a Vernier line-discrimination task eventually showed acuity greater than that of individual photoreceptors (Poggio, Fahle, & Edelman, 1992). Perceptual learning is implicated in these domains because the effects are evident at early stages of processing and are often restricted to the trained stimulus orientation (Notman et al., 2005), or even to a specific retinal location (Fahle, Edelman, & Poggio, 1995).

However, this does not appear to be the case where color categorization is concerned. In the case of color, humans may already have hyper-acuity (Chuchland & Sejnowski, 1994), so that no further 'tuning' occurs with category learning. Indeed, thresholds for color discrimination have been reported to be remarkably robust in some patient populations, despite a variety of other visual deficits including reduced sensitivity to flicker and reduced contrast sensitivity (Regan, Freudenthaler, Kolle, Mollon, & Paulus, 1998). With regard to the color domain, we believe that "categorical perception" is categorical but not perceptual, and should rather be referred to simply as a category effect.

Acknowledgements

This research is supported by Economic and Social Research Council Grant RES/221527 to the first author and by a Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2006-352-H00003) to Dr. Hyensou Pak. Correspondence concerning this article should be addressed to: Prof. Debi Roberson, Dept. of Psychology, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ. Tel: (+44) 01206 87-4175/3710, fax: (+44) 01206 873590, robedd@essex.ac.uk.

Appendix. C.I.E. u^*v^* values for the 15 stimuli for which thresholds were measured. *L* was held constant at 36.708.

HUE	Value	CHROMA	u [*]	v*
10.00G	5.47	7.03	0.131	0.434
1.25BG	5.47	7.03	0.132	0.472
2.50BG	5.47	7.03	0.133	0.470
3.75BG	5.47	7.03	0.135	0.465
5.00BG	5.47	7.03	0.138	0.462
6.25BG	5.47	7.03	0.140	0.460
8.50BG	5.47	7.03	0.142	0.458
10.00BG	5.47	7.03	0.143	0.455
1.25B	5.47	7.03	0.145	0.453
2.50B	5.47	7.03	0.147	0.451
3.75B	5.47	7.03	0.149	0.449
5.00B	5.47	7.03	0.151	0.446
6.25B	5.47	7.03	0.153	0.443
7.50B	5.47	7.03	0.156	0.441
8.75B	5.47	7.03	0.158	0.439

References

- Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition*, *12*, 83–95.
- Bornstein, M. H., & Korda, N. O. (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research*, 46, 207–222.
- Chuchland, P. S., & Sejnowski, T. J. (1994). *The computational brain.* Cambridge, MA: MIT Press.
- Davidoff, J., Davies, I., & Roberson, D. (1999). Color categories of a stoneage tribe. Nature, 398, 203–204.
- Drivonikou, G. V., Kay, P., Regier, T., Ivry, R. B., Gilbert, A. L., Franklin, A., et al. (2007). Further evidence that Whorfian effects are stronger in the right visual field than in the left. *Proceedings of the National Academy of Sciences of the USA*, 104, 1097–1102.
- Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. Vision Research, 35, 3003–3013.
- Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Current Biology*, 6, 292–297.
- Fletcher, R. (1980). The City University Colour Vision Test. Windsor, Berks: Keeler.

- Fonteneau, E., & Davidoff, J. (2007). Neural correlates of colour categories. Neuroreport, 18(13), 1323–1327.
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. Proceedings of the National Academy of Sciences of the USA, 103(2), 489–494.
- Goldstone, R. L. (1998). Perceptual Learning. Annual Review of Psychology, 49, 585–612.
- Harnad, S. (Ed.). (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. Cambridge: Cambridge University Press.
- Ishihara, S. (1992). Ishihara's tests for colour-blindness. Tokyo: Kanehara & Co. Ltd.
- Kay, P., Berlin B. & Merrifield, W. (1991). Biocultural implications of systems of color naming. *Journal of Linguistic. Anthropology*, 1, 12–25.
- Kim, Y. S., Pak, H., & Lee, Y. H. (2001). A study on Munsell color space for Korean color names. *Journal of Korean Society of Color Studies*, 15, 29–36.
- King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research*, 34, 885–912.
- Lee, M. Y., Kim, Y. S., Oh, K. K., Pak, H., Koo, M. M., Yeo, M. S., et al. (2003). The standardization study of the names of non-luminous object colours and light-source colours (Research report). Korean Agency for Technology and Standards.
- Notman, L. A., Sowden, P. T., & Ozgen, E. (2005). The nature of learned categorical perception effects: A psychophysical approach. *Cognition*, 95, B1–B14.
- Özgen, E. D., & Davies, I. R. L. (2002). Acquisition of categorical color perception: A perceptual learning approach to the linguistic relativity hypothesis. *Journal of Experimental Psychology: General*, 131(4), 477–493.
- Pilling, M., Wiggett, A., Özgen, E., & Davies, I. R. L. (2003). Is color "categorical perception" really perceptual? *Memory and Cognition*, 31, 538–551.
- Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. Science, 256, 1018–1021.
- Regan, B. C., Freudenthaler, N., Kolle, R., Mollon, J. D., & Paulus, W. (1998). Discrimination thresholds in Parkinson's disease: Results obtained with a rapid computer-controlled color vision test. *Vision Research*, 38, 3427–3431.
- Roberson, D., Davidoff, J., & Braisby, N. (1999). Similarity and categorisation: Neuropsychological evidence for a dissociation in explicit categorisation tasks. *Cognition*, 71, 1–42.
- Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory and Cognition*, 28, 977–986.
- Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. (2005). Color categories in Himba: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, 50, 378–411.
- Roberson, D., & Hanley, J. R. (2007). Color categories vary with language after all. Current Biology, 17, 605–606.
- Roberson, D., Pak, H. S., & Hanley, J. R. (2008). Categorical perception of color in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, 107, 752–762.
- Tan, L.H., et al. (2008). Language affects patterns of brain activation associated with perceptual decision. Proceedings of the National Academy of Sciences of the USA, 105, 4004-4009.
- Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J-R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences of the USA*, 106, 4567–4570.
- Wiggett, A. J. D., & Davies, I. R. L. (2008). The effect of Stroop interference on the categorical perception of color. *Memory and Cognition*, 36, 231–239.
- Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. Proceedings of the National Academy of Sciences of the USA, 104(19), 7780–7785.
- Zohary, E., Celbrini, S., Britten, K. H., & Newsome, W. T. (1994). Plasticity that underlies improvement in perceptual performance. *Science*, 263, 1289–1292.