DT2350 Human Perception for Information Technology

Multisensory processing: an introduction

Roberto Bresin

Literature

Charles Spence (2011) **Crossmodal correspondences: A tutorial review**. *Atten Percept Psychophys*, 73: 971-995

Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People.

Chapter #100*:* ***People value a product more highly when it's physically in front of them***

Why do we have a brain?

Sea squirt

Crossmodal correspondences

Charles Spence (2011) Crossmodal correspondences: A tutorial review

How does the brain “know” which stimuli to combine?

➔ Research show that people exhibit consistent crossmodal correspondences between many stimulus features in different sensory modalities.

Example:
high-pitched sounds → small & bright objects, high up in the space

Multisensory perception

Cognitive neuroscience research on the topic of multisensory perception focuses on trying **to understand**, and **to model** the **spatial and temporal factors modulating multisensory integration.**

Temporal factor: Multisensory integration is more likely to occur the closer that the stimuli in different modalities are presented in time.

Spatial factor: Spatial coincidence facilitates multisensory integration.

Other factors influence multisensory integration in humans

Semantic congruency: refers to situations in which pairs of auditory and visual stimuli are presented that vary (i.e., match vs. mismatch) in terms of their identity and/or meaning.

Synaesthetic congruency: refers to correspondences between more basic stimulus features (e.g., pitch, lightness, brightness, size) in different modalities.

➔ **Unity effect**: stimuli that are either semantically or synaesthetically congruent will more likely be bound together

Synaesthetic congruency ➔ Crossmodal correspondences

Crossmodal correspondence

Def: *Compatibility effect between attributes or dimensions of a stimulus (i.e., an object or event) in different sensory modalities (be they redundant or not).*

Crossmodal correspondences occur between polarized stimulus dimensions: a more-or-less extreme stimulus on a given dimension should be compatible with a more-or-less extreme value on the corresponding dimension.

Key feature: crossmodal correspondences are **shared by a large number of people**.

Crossmodal associations

Example: sound symbolism

Mil

Mal

Sapir (1929)

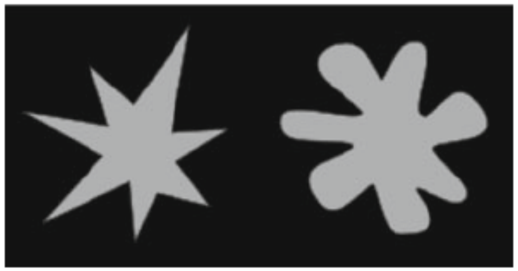
Maluma

Takete

Köhler(1929)

bouba

kiki



Ramachandran & Hubbard (2001, 2003)

Crossmodal associations

Example: sound symbolism

The bouba/kiki effect does not show in:

- **Children with autism spectrum disorde**r
- **People with damage to the angular gyrus** (located within the temporal–parietal–occipital [TPO] region)

➔ This suggests that **crossmodal correspondences** (at least those involving sound symbolism) **can occur at quite a high leve**l.

Crossmodal associations: across groups

Adults and children (5-year-olds) reliably matched brightness with loudness crossmodally:

Light grey patches → Louder sounds

Darker grey patches → Softer sounds

Crossmodal associations

Unidimensional sensory stimuli

high-pitched tones → brighter surfaces

louder sounds → visual stimuli with higher contrast

Complex stimuli

music → pictures

Crossmodal associations

20-30 day-old
Loudness → Brightness

Infants (3-4 month-old)
Pitch → Visual elevation

2 year-old children
Loud sounds → Large shapes

➔ The ability to match **other dimensions** crossmodally appears to **develop** somewhat **more slowly**

Crossmodal associations: other than auditory/visual stimuli

Vision → Touch

Audition → Touch

Taste/Flavours → Sounds

Colour → Odours

Colour → Tastes

Colour → Flavours

Pitch → Smell

Smells → Shapes

Shapes → Tastes/Flavours

➔ Likely that crossmodal correspondences exist between all possible pairings of sensory modalities.

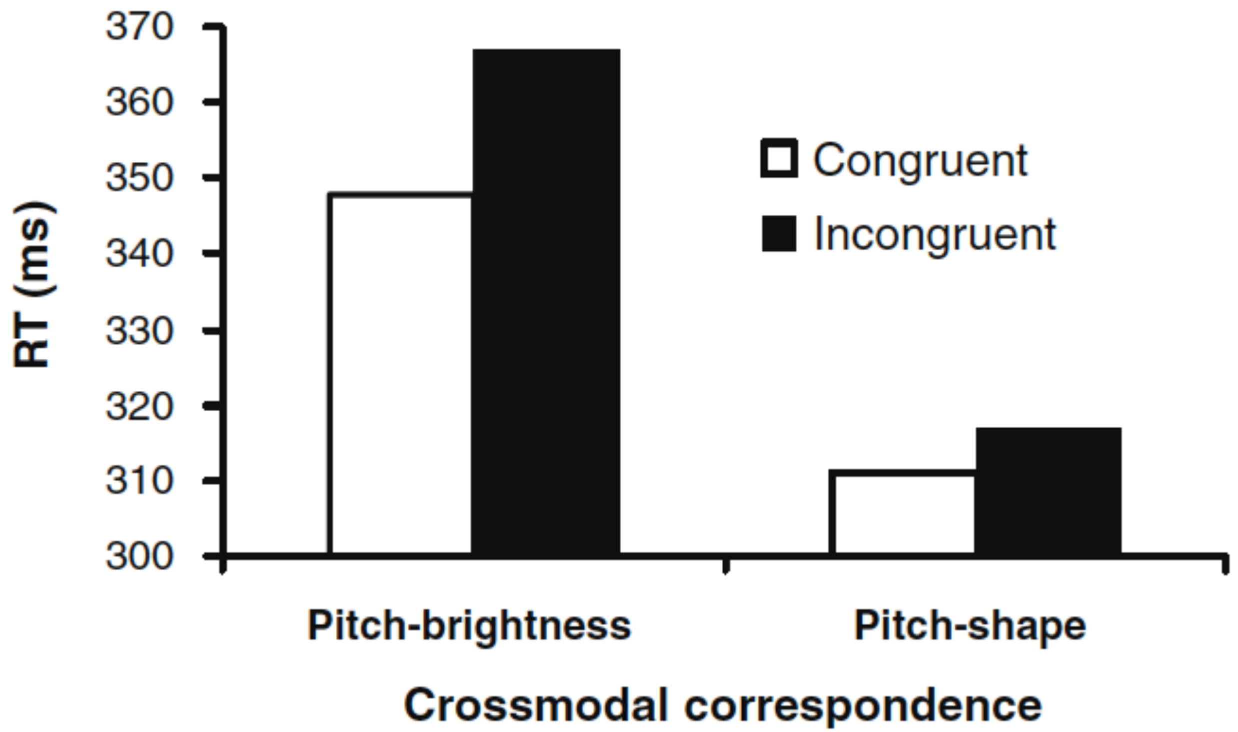
Assessing the impact of crossmodal correspondences

Consolidated correspondence:

Pitch (100 or 1000 Hz) → Elevation (upper or lower corner)

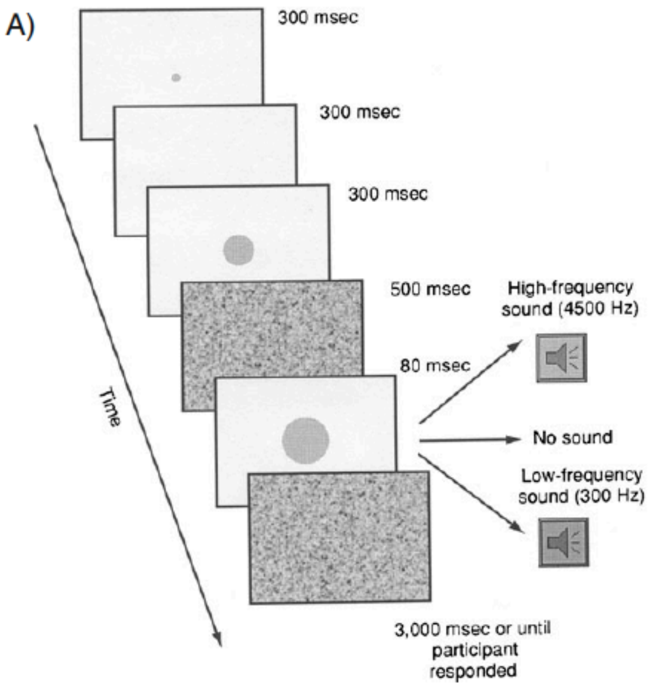
Harder to classify the size of a visual stimulus (as either large or small**) when the task irrelevant sound presented on each trial is incongruent in pitch** (e.g., when a high-pitched tone is presented at the same time as a large target) **than when the distractor sound is congruent** (e.g., when a low tone is presented with the large target).

Response time (RT)



Task irrelevant sound (of either relatively low or high pitch) can significantly influence participants' responses on a speeded visual size discrimination task

A sound (300 or 4500 Hz) was presented in synchrony with the second disk (otherwise, no sound was presented)



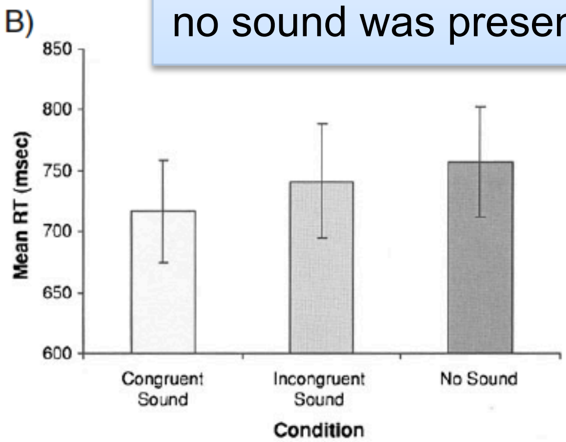


Fig. 3 **a** Schematic time line and stimuli from Gallace and Spence's (2006) experiments highlighting the consequences of the crossmodal correspondence between auditory pitch and visual size for participants' speeded discrimination responses. **b** Results from the speeded "same vs. different" visual size discrimination task. Congruent pairs of stimuli (e.g., a larger disk paired with the lower tone) gave rise to faster RTs. The error bars indicate the standard errors of the means. [From Figs. 1 and 5 of "Multisensory Synesthetic Interactions in the Speeded Classification of Visual Size," by A. Gallace and C. Spence, 2006, *Perception & Psychophysics*, *68*, pp. 1191–1203. Copyright 2006 by the Psychonomic Society. Redrawn with permission.]

Crossmodal correspondences that have been shown to influence participants' RT

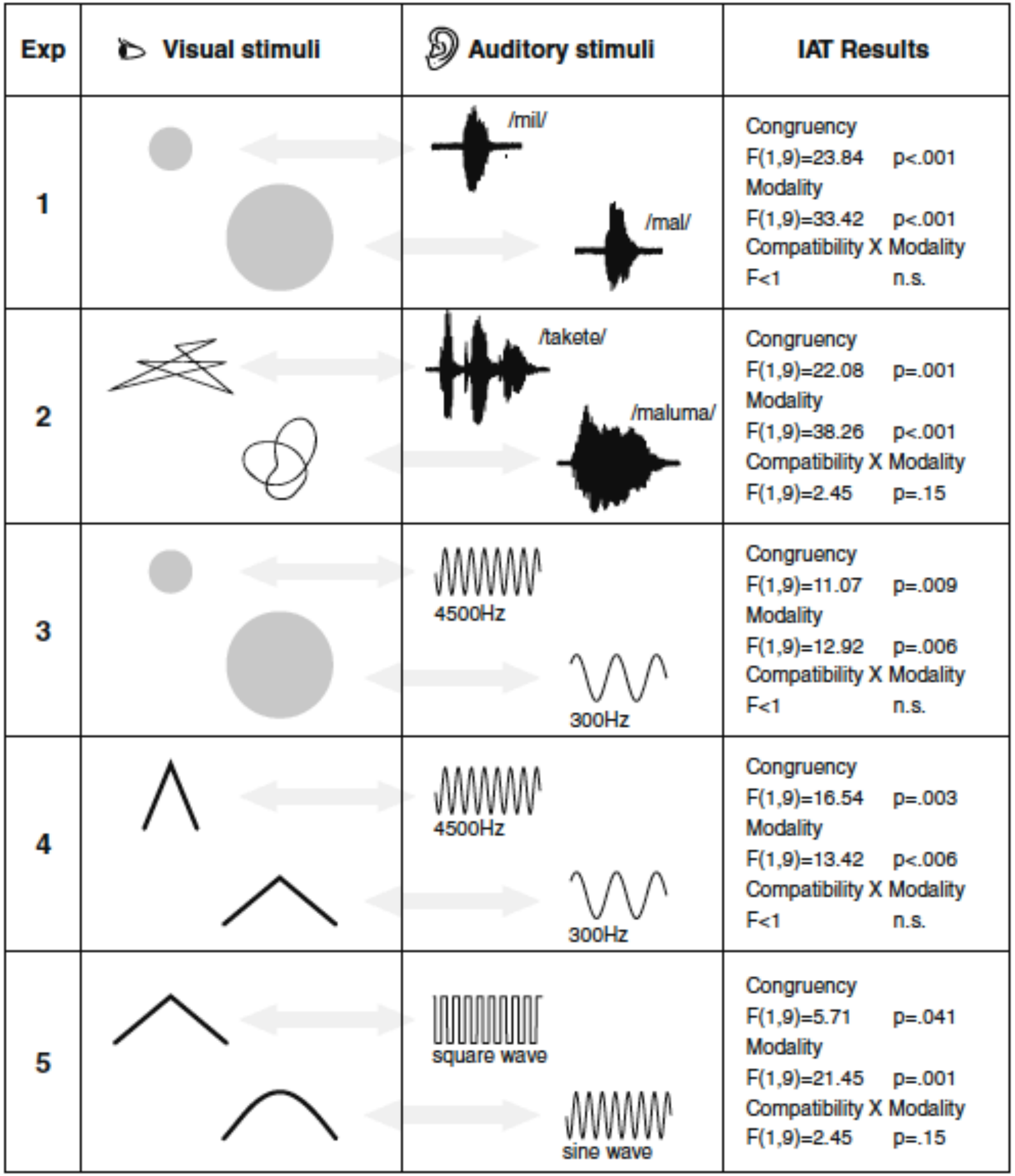
Table 1 Summary of crossmodal correspondences that have been shown to influence participants' performance on the speeded classification task together with the null results that have been reported to date

Auditory Dimension	Visual Dimension	Crossmodal Correspondence	High-Pitch/Loud Sound Corresponds to:	Studies
Pitch	Elevation	Yes	High elevation	Ben-Artzi and Marks (1995); Bernstein and Edelstein (1971); Evans and Treisman (2010); Melara and O'Brien (1987); Patching and Quinlan (2002)
	Brightness	Yes	Brighter stimulus	Marks (1987a)
	Lightness	Yes	Lighter stimulus	Marks (1987a); Martino and Marks (1999); Melara (1989a)
	Shape/angularity	Yes	More angular shape	Marks (1987a)
	Size	Yes	Smaller object	Evans and Treisman (2010); Gallace and Spence (2006)
	Spatial frequency	Yes	High spatial frequency	Evans and Treisman (2010)
	Direction of movement	Yes	Upward movement	Clark and Brownell (1976)
	Contrast	No	N/A	Evans and Treisman (2010)
	Hue	No	N/A	Bernstein, Eason, and Schurman (1971)
Loudness	Brightness	Yes	Brighter stimulus	Marks (1987a)
	Lightness	No	N/A	Marks (1987a)

Note. Other crossmodal correspondences demonstrated using other tasks are not mentioned here on the speeded classification task

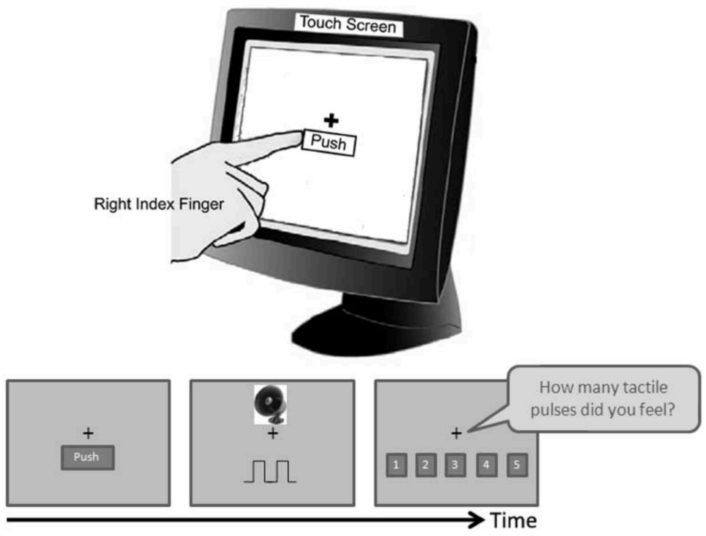
Audiovisual crossmodal correspondenses and sound symbolism

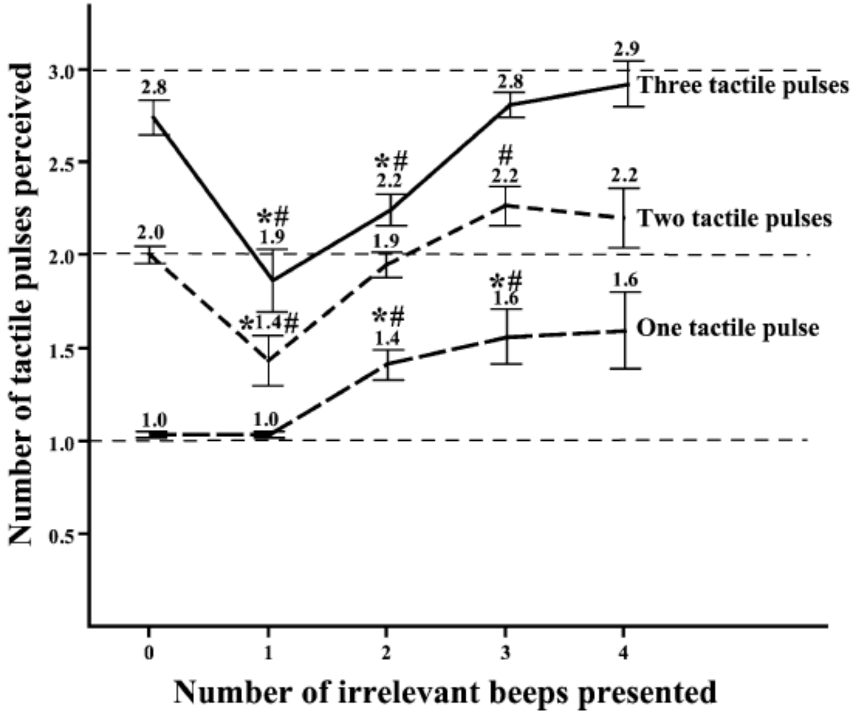
(Parise & Spence 2012)



Feeling what you hear: task-irrelevant sounds modulate tactile perception delivered via a touch screen

Lee & Spence (2008)





The music of taste

Knöferle & Spence (2012) Crossmodal correspondences between sounds and tastes

Table 1 Summary of crossmodal correspondences between basic tastes and sonic elements demonstrated to date

Author(s)	Auditory property	Sweet	Sour	Salty	Bitter
Bronner, 2012	Sharpness/spectral balance	low	high		
	Roughness	low	high		
	Ambitus	small	large		
	Articulation	legato	staccato		
	Rhythm	even	syncopated		
	Melodic intervals	small	large		
	Melodic consonance	consonant	dissonant		
	Tempo	slow	fast		
Crisinel & Spence, 2009	Pitch		high		low
Crisinel & Spence, 2010a	Pitch	high	high	average	low
	Instrument type	piano	brass	brass	brass
Crisinel & Spence, 2010b	Pitch	high	high	ns	ns
Crisinel & Spence, 2012	Pitch	higher			lower
	Instrument type	piano			ns
Knöferle & Spence, 2012	Pitch	high	average	average	low
	Roughness	low	high	average	high
	Sharpness/spectral balance	ns	high	ns	low
	Discontinuity	low	high	high	high
	Attack	ns	ns	ns	ns
	Speed	ns	fast	ns	slow
Mesz et al., 2011	Pitch	average	high	low	low
	Articulation	legato	average	staccato	legato
	Loudness	soft	average	average	average
	Chord consonance	consonant	dissonant	average	average
	Melody consonance	consonant	dissonant	average	average
Ngo et al., 2011[2]	Consonants	soft			hard
	Vowel backness	back			front
Simner et al., 2010	Vowel height	higher	lower	lower	lower
	Vowel backness[3]	back	front	front	front
	Discontinuity	lower	higher	ns	higher
	Spectral balance	lower	higher	ns	ns

Sonic food: sound for crispy chips

Ig NOBEL NUTRITION PRIZE 2008.
Massimiliano Zampini of the University of Trento, Italy and Charles Spence of Oxford University, UK, **for electronically modifying the sound of a potato chip to make the person chewing the chip believe it to be crisper and fresher than it really is**.
http://www.improbable.com/ig/winners/#ig2008

http://news.bbc.co.uk/2/hi/science/nature/7650103.stm

(4 min interview with Charles Spence, including sound examples)

More examples

Hearing temperature of a drink: Hot/Cold pouring sound

http://www.npr.org/2014/07/05/328842704/what-does-cold-sound-like

Music manipulates taste/choice: French vs German wine

North et al. (1999) The influence of in-store music on wine selections. J. of Applied Psychology

Links

Crossmodal processes

https://www.facebook.com/xmodal

Multisensory Perception and Action

http://www.uni-bielefeld.de/(en)/biologie/cns/index.html

The Social Mind and Body Group (SOMBY)

http://somby.info

People value a product more highly when it's physically in front of them

Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People. Chapter #100

Example: bidding for a product (food or toys)

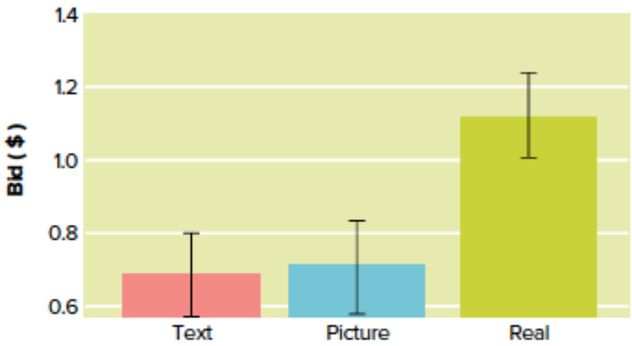


FIGURE 100.1 People valued the food more when it was in front of them

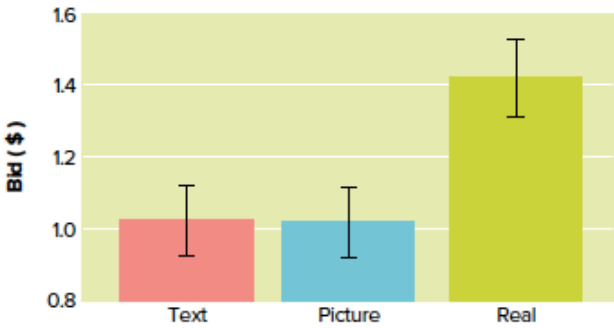


FIGURE 100.2 People valued the toys and trinkets more when they were physically present

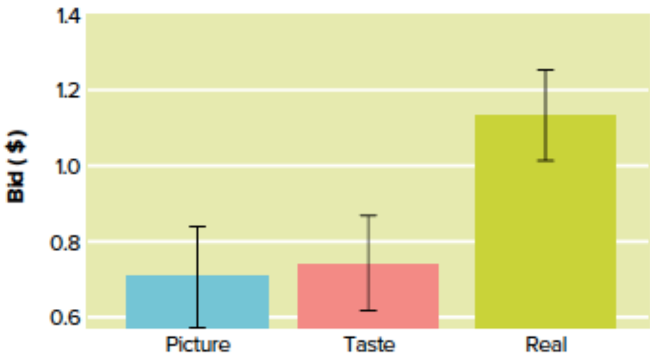


FIGURE 100.3 Samples (taste) were less effective than the actual product

People value a product more highly when it’s physically in front of them

Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People. Chapter #100

A PAVLOVIAN RESPONSE?

Bushong and his team hypothesize that there’s a Pavlovian response going on: when the product is actually available, it acts as a conditioned stimulus and elicits a response. Images and even text could potentially become a conditioned stimulus and produce the same response, but they have not been set up in the brain to trigger the same response as the actual item.

Takeaways

* Brick-and-mortar stores may retain an edge if they have products on hand, especially when it comes to price.
* Having a product behind glass or any other kind of barrier may lower the price that the customer is willing to pay.

The end

FINE