
The Instant Impact of Onscreen Aesthetics: The Effects of Typeface Personality

Richard L. Hazlett

Johns Hopkins University
School of Medicine
600 N. Wolfe St.
Baltimore, MD 21287 USA
rlhazlet@jhmi.edu

Kevin Larson

Microsoft Advanced Reading
Technologies
Microsoft Corp
One Microsoft Way
Redmond, WA 98052 USA
kevlar@microsoft.com

A. Dawn Shaikh

Software Usability Research
Laboratory
Wichita State University
Wichita, KS 67260 USA
adshaikh@gmail.com

Barbara S. Chaparro

Software Usability Research
Laboratory
Wichita State University
Wichita, KS 67260 USA
barbara.chaparro@wichita.edu

Abstract

This paper describes two studies that investigate the effects of congruency in the personality of typeface with onscreen content. In study 1 trait words were flashed onscreen in either trait congruent or incongruent typeface. Congruent typeface enhanced affective word classification as evidenced by shorter response latencies. Study 2 briefly flashed pages onscreen in varying fonts and found that participants were influenced by the font personality in their judgments about the emotional tone of the page. It is concluded that aesthetics such as typeface can influence reader's emotional judgments about onscreen content. These methods appear useful for evaluating onscreen aesthetics.

Keywords

Affect, typography, design, aesthetics, research methods, reading.

ACM Classification Keywords

H.5.2: User Interfaces-Evaluation/Methodology, Screen design

Copyright is held by the author/owner(s).

CHI 2008, April 5 – April 10, 2008, Florence, Italy

ACM 978-1-60558-012-8/08/04.

Introduction

There is an abundance of evidence that people can develop preferences and be affected emotionally with just very brief visual presentations [e.g., 20,24]. For example, Willis and Todorov [23] found that participants were able to judge the traits of people after viewing their face for only 100ms. Increasing the exposure time did not change the trait judgments. More exposure time allowed for more finely differentiated judgments, but later judgments were already anchored on the initial impression.

A number of authors have characterized humans as having two evaluative systems [3]. The first system is automatic, fast and preconscious. The second system involves conscious thinking, and is slow, effortful and deliberate. In his influential 1980 article, Zajonc [24] illustrated how people can develop preferences based on brief glimpses without conscious awareness. Furthermore people are often unable to explain what in the environment has influenced their preference or behavior [10]. Many researchers have used priming and other implicit methods to demonstrate that the automatic fast system has a lasting effect on people's stated preferences, deliberate thought, and actions [4, 5, 16].

Onscreen page aesthetics are processed quickly and automatically as well, and can impact readers with very brief presentations. Lindgaard, Fernandes, Dudek, and Brown [14] found that in just 50 ms users make aesthetic judgments on their preference for a web site. Internet site pages were flashed onto a computer screen and respondents made a decision about how appealing they found the site. Their preferences were similar to the ones they made at longer exposures, and

the impression was lasting and influenced the rest of their experience with the Internet site. Lindgaard et al. [14] were not able to determine in their study what common aesthetic features their participants found most appealing. They [14] did conclude that these first impression judgments based on brief glimpses are more the result of an emotional, physiologic response and before cognitive appraisal had had time to occur. Some visual information such as hue, orientation, size and motion can be processed from just a single glimpse, and this has been termed pre-attentive processing [9, 20, 21]. It occurs before visual search has been initiated. Impressions based on such brief glimpses Zajonc labeled as the 'mere exposure effect' [24]. The 'wow' response, a burst of positive emotion when first viewing a product or feature, is a particular type of first impression reaction that has been found to have a measurable physiologic component [8].

A particular onscreen visual aesthetic that is so ever-present that it is usually not given a conscious thought is typeface. Though typeface may not be the focus of a reader's attention, it is often credited with creating a page or document's first impression. Typefaces are typically classified according to unique typographic features (serif, bold, condensed, etc) and overall appearance (humanist, modern, old style, etc.) [12]. The combination of appearance and typographical features often lead graphic artists and typographers to describe typefaces using personality traits [1].

Typographers and designers are interested in the typeface personality or "typographic allusion" which refers to "the capacity of a typestyle to connote meaning over and above the primary meaning which is linguistically conveyed by words" ([13] p. 243).

Brumberger [2] describes the Bauhaus school of design and their belief that the “content and purpose of the text should dictate the design – the form – of a document, and that form, including typography, should express the content just as the verbal text itself expresses content” (p. 207). Within communications research, many experts suggest that typefaces can convey mood, attitude, and tone while having a distinct persona based on the font’s unique features. Each document should be rendered in a font that connects the mood, purpose, intended audience, and context of the document.

Most empirical research concerning typefaces focuses on the legibility or readability with little concern for the perceived personality. One notable exception to this was a study on the effects of perceptual qualities of print typeface by Lewis and Walker [13]. Using a Stroop-like paradigm participants saw adjectives (e.g., fast/slow, heavy/light) descriptive of one of two print typefaces through a tachistoscope and had to press certain keys based on which adjective word was being displayed. Reaction times were slower for typefaces that were incongruent with the adjective being tested. Based on research on visual processing [18,20] Lewis and Walker [13] speculated that the activation of semantic categories associated with typeface qualities proceeds more rapidly and is available sooner than the activation of semantic categories associated with word meaning. The visual aesthetic information had a quick effect that influenced the latter processing of word meaning.

Reaction time measures like that used by Lewis and Walker [13] have been employed in a variety of paradigms to investigate emotional processing and

automatic activation of attitudes. The brief presentation of emotion words flashed onscreen has been shown to affect subsequent lexical decisions [4,5]. In the Stroop matching task Goldfarb and Henik [6] found that congruent word meaning/color combinations were processed faster than incongruent. This Stroop facilitation effect is particularly relevant for studying congruency effects of typeface personality. Unlike priming methodologies, where the facilitating prime is presented before the target word, the facilitating colorword in the Stroop test and the facilitating typeface personality are presented simultaneously with the content word [15, 18].

Onscreen fonts or typefaces have been found to be perceived as having certain personality traits. Using an online survey method Shaikh, Chaparro and Fox [19] had 561 participants rate 20 fonts using 15 adjective pairs. The adjective pairs included pairs like stable/unstable, conformist/rebel, sad/happy. These adjectives were more human-like, affectively toned, and at a higher level of abstraction than the perceptual quality adjectives of fast/slow, heavy/light used by Lewis and Walker [13]. Factor analysis found that personality traits were attributed to fonts based on their design family, and clustered into five factors (Serif, Sans-Serif, Modern, Monospace, Script/Funny).

The impact of onscreen typography and other reading aesthetics has been investigated with reading performance measures like speed and comprehension [11]. Gugerty, Tyrrell, Aten, & Edmonds [7] found reading speed and comprehension advantages for the advanced ClearType rendering engine over the basic black & white rendering engine. In a lexical decision task they also found greater word recognition for words

in ClearType. But other more subtle on-screen aesthetic improvements don't demonstrate the large performance differences that we see with ClearType. Larson, Hazlett, Chaparro and Picard [11] investigated the performance difference between documents with good page layout and poor page layout and found no speed or comprehension differences between these two conditions. Investigating the effects of on-screen aesthetics may require a different approach.

Congruent Corsiva	Congruent TNR	Incongruent Corsiva	Incongruent TNR
<i>pretty</i>	reliable	<i>stable</i>	lithe
<i>gorgeous</i>	cold	<i>inhibited</i>	lovely
<i>stylish</i>	calm	<i>conformist</i>	charming
<i>trendy</i>	staid	<i>aloof</i>	moody
<i>sexy</i>	constant	<i>proper</i>	attractive

figure 1. Congruent and incongruent presentations of trait words in Times New Roman (TNR) and Corsiva fonts.

Experimental Objectives

We wanted to continue the line of research begun by Shaikh et al [19] and investigate if these more human-like personality traits of fonts could impact on the reading experience and affect the communication of written material. Our hypotheses were that congruent content and typeface pairings, as opposed to incongruent content and typeface pairings, enhance and facilitate the communication of the emotional tone of the content. Typefaces that have a congruent emotional tone contribute to the activation of the emotion category of the content, enhancing the speed

of processing emotional meaning of content, and deepening the activation. Secondly, we wanted to investigate whether the personality of the typeface will influence the perception of the emotional tone of an onscreen page. Thirdly, we are interested in whether the priming and brief exposure methods used in these studies could be useful in HCI investigations of the effects of onscreen aesthetics.

Study 1: Word test of the effects of font personality

Method

Twenty-five participants read emotionally congruent and incongruent word/font pairings quickly flashed on a computer screen. Participants were asked to decide whether it was a positive or negative emotional word by clicking a button on the screen. The study sample consisted of mainly hospital employees or students between the ages of 21 to 40 years. The inclusion criteria were a minimum of one year of college, either corrected or uncorrected 20/20 vision, and participants read from a computer screen 5 or more hours per week.

Times New Roman (TNR) typeface was compared to Monotype Corsiva. These fonts were chosen based on the results of the study on personality of fonts by Shaikh et al. [19]. These two fonts were found to have extremely different personality traits. TNR's distinctive trait words were stable, conformist, mature, and practical, while Corsiva's trait words were feminine, elegant, and attractive. A list was developed with 48 words congruent with TNR's traits, and 48 words congruent with Corsiva's traits. The list was developed by using a thesaurus and also selecting words from a

list of trait words that shared the same personality factor as the font's trait [22]. Twenty-four nonrelated trait words were mixed in the presented list to make a pseudo-random presentation of 120 words. Twelve practice words were included at the beginning for a total of 132 trials. A mixture of positive and negative trait words was presented. The words were presented in TNR, Corsiva, and Verdana fonts. Half of the Corsiva and TNR trait words were presented in TNR font, and half in Corsiva font, and the nonrelated words were presented in Verdana. The trait words were counterbalanced between TNR and Corsiva. Examples of congruent and incongruent trait words presented in TNR and Corsiva are presented in Figure 1.

Font Presentation			
Traits	Congruent	Incongruent	Difference
Positive	1276 (240)	1320 (288)	44
Negative	1632 (519)	1635 (524)	3

table 1. Mean (SD) response latency in ms to trait words.

Participants saw the word appear in the middle of the computer screen, and used a mouse to click on either the positive or negative button to indicate whether the word was a positive or negative trait. The word would disappear after the button was clicked. 4 seconds later a new trial would start with a marker rectangle. The new word would then appear 500 ms after the rectangle. The selection and the elapsed time between presentation of word and button click was recorded. The participants used a Dell Inspiron 600m laptop with a 14.1 inch screen at a resolution of 1400 x 1050 pixels

(124dpi). The target word was presented in 14 point font.

Results

Reaction time responses 2.5 standard deviations above the participant's mean were considered outliers and eliminated. Nine of the Corsiva words and 12 of the TNR words were endorsed as negative by more than half the participants. The mean response latency in milliseconds for the positive and negative trait words are presented by congruency and valence in Table 1.

A 2 (Congruency) x 2 (Valence) repeated measure MANOVA using the natural log of the response time (to normalize the data) found the congruency effect to approach significance [$F(1,24) = 3.84, p = .062$], and the valence effect to be significant [$F(1,24) = 88.08, p < .000$]. A follow up paired samples *t*-test of the positive words found that the congruent font presentations had a quicker response time than the incongruent presentations [$t(24) = 2.30, p = .031$]. A paired samples *t*-test of the negative words found no difference between congruent and incongruent presentations [$t(24) = -0.09, p = .93$]. These results indicate that in making decisions about the emotional tone of positive words, a congruent font facilitates decision making, while decision making for negatively toned words did indeed take significantly longer than making decisions about positive words, and wasn't influenced by font.

Study 2: Page test of the effects of font personality

This study tested font personality effects at page level and compared the fonts Georgia, Arial and Corsiva. The serif typeface Georgia and the sans serif typeface Arial

were selected based on data from Shaikh et al. [19], and provided a more challenging test of the font congruency effect by testing two typefaces that are commonly used in text documents.

Method

A different set of 24 participants with similar demographics were tested in this study. The participants were briefly presented a page of several paragraphs in one of three typefaces. Participants saw for 700ms onscreen presentations of one of three typefaces in 10-point font: Arial, Georgia, or Corsiva. After the page disappeared from the screen, participants were asked to rate how well (on a 4-point scale) the tone of the page matched one of six trait words, with 4 being very much like the trait word, and 1 very much like the antonym of the trait word. The response scale was anchored with the trait word and its antonym. The top three traits for Georgia (practical, formal, and assertive) and Arial (stable, conformist, and unimaginative) were selected based on the Shaikh et al. data (see [19] for more detail on how the font trait rankings were developed).

The 700ms presentation exposure was selected based on pilot testing results. This exposure time was found to give participants enough of a glimpse of the page to make a judgment on the 4-point scale without being able to read more than 3 or 4 words. Time to read several words insured that the font was processed. Participants were asked to make their selection based on the 'way the page looked' and not the content.

Participants were questioned on each of the six trait words twice for each of the three page font presentations. This resulted in 36 experimental trials

and an additional 4 practice trials at the beginning. There were six pages presented that had the exact same sentences but with sentence order and paragraph breaks varied thus creating the perception of varying content. The order and pairing of the passages and fonts were counter balanced and randomly varied.

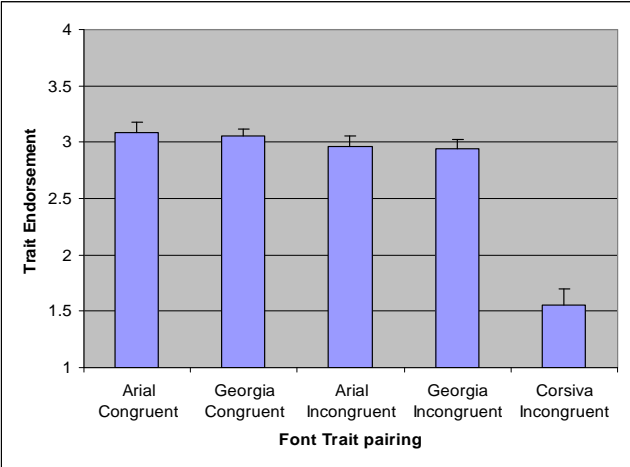


figure 2. Font congruency effects for brief page glimpse. Error bars represent the standard error of the mean. Responses ranged from 1-4, with 4 most like the trait.

Results

Mean responses were calculated for trait choices made with pages presented in each of the five conditions: Congruent Arial, Congruent Georgia, Incongruent Arial, Incongruent Georgia, and Incongruent Corsiva (figure 2). Georgia is congruent with the trait words practical, formal, and assertive, while Arial is congruent with the trait words stable, conformist, and unimaginative. The Corsiva font presentations were always classified as

<i>Trait</i>	<i>Rank Order^a Difference</i>	<i>Rating Scale^a Difference</i>	<i>Congruent Font Rating^b</i>	<i>Incongruent Font Rating^b</i>	<i>Congruent/Incongruent Difference</i>
Unimaginative	6	.16	2.14 (.85)	1.67 (.87)	.47
Assertive	5	.13	1.93 (.58)	1.50 (.38)	.43
Practical	5	.09	2.17 (.55)	2.02 (.79)	.15
Formal	4	.25	2.21 (.57)	2.23 (.40)	-.02
Conformist	2	.03	1.81 (.74)	1.86 (.74)	-.05
Stable	2	.07	2.24 (.70)	2.52 (.51)	-.28

table 2. Page glimpse mean (SD) scale scores by congruency and trait. Notes. ^a The difference between the Arial and Georgia font ratings are from Shaikh et al [19]. ^b Trait rating is on a 1-4 scale, with 4 being the highest.

incongruent, because there were no Corsiva trait questions. Two participants had invalid data due to software problems, leaving 22 valid participants.

A repeated measure ANOVA found a significant linear effect [$F(2,20) = 31.09, p < .001$] for congruency condition. Follow up tests comparing the conditions found that the Corsiva incongruent condition was significantly less than both the Arial/Georgia congruent and incongruent conditions [$p < .001$]. The overall mean difference between the congruent and incongruent Arial/Georgia conditions trended to congruent being more strongly endorsed for the trait [$p = .10$].

The Arial and Georgia effects were the average across all 6 traits, and these trait words varied considerably on how far apart their ranking was for the two fonts. Therefore, each of the six traits was compared separately for font congruency effects between the Arial and Georgia fonts. However, the ranks were definitely not equal across traits – for some, the top 5 or 6 were

very close to one another in scale score, while for others the top 5 or 6 were quite disparate. This variability is seen in the rating scale differences between fonts listed in Table 2.

In Table 2 the rank and scale differences from Shaikh et al [19] are reproduced and the traits are listed in descending rank order. The means (SD) for congruent and incongruent brief page glimpse answers for the Arial and Georgia fonts are added. Positive congruency difference values indicate that the congruent pairing was more strongly endorsed for the trait than the incongruent pairing, and negative values indicate the incongruent pairing was more strongly endorsed.

Such a small sample of traits precludes statistical analysis, but there are some trends of interest to be seen from the table. The congruency difference for the Assertiveness and Unimaginative traits are the largest, and these two traits also have the most spacing between rankings for the two typefaces. Those traits appear to differentiate the two typefaces the most, with

Georgia being most assertive and Arial most unimaginative.

These results suggest that the personality of a font or typeface can influence a reader's perception of the emotional tone of an onscreen page of text. Also this effect can occur during a brief glimpse of the page before the reader has the opportunity to read more than several words.

Overall Discussion

The results of study 1 indicate that a typeface congruent with the word's emotional meaning can facilitate the affective classification of that word. This result was only true for words with a positive emotional valence. The difference in elapsed time of 44ms may seem small, but it is inline with the effects commonly found in the affective priming literature [15]. More research will be needed to understand what influence valence has on the effects of font personality, or more generally, onscreen aesthetics.

Study 2 demonstrated that a brief glimpse of an onscreen page of text can communicate emotional tone based on the personality of the typeface. This methodology appears useful because the processing of page content is minimized so the effects of the aesthetics of the page are more clearly studied. Specifically, the typeface of the words on the page, even though no more than several of the words were read, influenced the rating of the emotional tone of the page. The congruency effect was more clearly noted where there were larger differences between trait words and typeface personality.

There was some suggestive evidence in these studies that the congruency effect was stronger for traits that had the greatest ranking difference between fonts. Typeface closer in personality like Georgia and Arial may only influence people on one or two signature traits. In this study there were too few traits to really describe and test this relationship, and future studies would need to test with more traits and a greater variance in font rank differences.

This study adds to the Lindgaard et al [14] findings in that it demonstrates that the initial response to visual aesthetics can be more complex personality evaluations and not just an overall rating on how appealing was the visual. Certainly though, how complex the evaluation can be is going to be a function of the exposure time, and what we found at 700ms likely does not apply to what is possible at 50ms. In fact, controlling exposure time could be a viable method for targeting the type of processing effects one is interested in. Identifying optimal exposure durations for studying different aspects of interfaces would be a useful line of research. One obvious variable is how much visual search is conducted before the display is evaluated.

When the differences in personality were slight, we found that we were reaching the limits of power for the methods. Larger samples and good precision in experimental design should help improve the power to test the more subtle differences. Subtle differences like that found between Georgia and Arial fonts are likely more similar to the practical questions that face design teams everyday, and trade offs between sample size and power become important choices. Therefore further research in this area that can identify improvements and enhancements to these types of methods would be

useful for the practical application of evaluating the influence of interface aesthetics. This study can only be considered an initial attempt at that task.

These research findings suggest that like other areas of humans interacting with their environment, human computer interactions are affected by fast automatic evaluations in response to subtle visual background aesthetics. Though this study did not investigate this point, other studies have found that these automatic evaluations and the data used to make them are often not available to consciousness [10, 17]. Sometimes no amount of introspection can discover the aspects of an environment that evoked a certain impression or preference, and verbal reports are only post-hoc rationalizations [17]. This quality of first impressions limits the validity of verbal reports or user questionnaires for discovering the influential features of the user environment, and points to the need for more implicit methods in understanding the effects of onscreen aesthetics.

A major interest in the effects of onscreen aesthetics would be determining what aspects of the visual display evoked a particular persona or appeal. Though in their pioneering study Lindgaard et al [14] were not successful in such a determination, these methods may yet hold promise for such an understanding, and future research could be designed with that purpose in mind.

A number of studies have demonstrated [e.g., 10] that contextual aesthetics are important factors influencing task behavior and perception, and this study only reaffirms that finding for onscreen visuals. Just how far reaching is the impact of onscreen aesthetics is a topic for future research. This study examined a very small

set of onscreen aesthetics, and certainly the conclusions drawn from this study are limited. At this point, it seems safe to suggest that congruency of typeface and other aesthetics with the onscreen content can only enhance communication in HCI.

Acknowledgements

This research was sponsored by a grant from the Advanced Reading Technologies Group at Microsoft Corporation.

References

- [1] Berry, J.D. *Now read this: The Microsoft ClearType font collection*. Microsoft Corporation, Seattle, WA, 2004.
- [2] Brumberger, E. The Rhetoric of Typography: The Persona of Typeface and Text. *Technical Communication*, 50(2), (2003), 206-223.
- [3] Chaiken, S., & Trope, Y. (Eds.). *Dual process theories in social psychology*. (1999), New York: Guilford Press.
- [4] Fazio, R. H. On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, 13, 2 (2001), 115-141.
- [5] Fazio, R. H., Sanbonmatsu, D.M., Powell, M.C., & Kardes, F.R. On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, (1986), 229-238.
- [6] Goldfarb, L., & Henik, A. New data analysis of the Stroop Matching Task calls for a reevaluation of theory. *Psychological Science*, 17, (2006), 96-100.
- [7] Gugerty, L., Tyrrell, R.A., Aten, T.R., & Edmonds, K.A. The effects of sub-pixel addressing on users' performance. *ACM Transactions on Applied Perception*, 1, (2004), 81-101.
- [8] Hazlett, R. L. and Benedek, J. Measuring the emotional reaction to passive first impression of

software. In *Proc. Designing Pleasurable Products and Interfaces 2005*, Eindhoven, Netherlands, Springer-Verlag.

- [9] Healey, C., Booth, K., & Enns, J. Visualizing real-time multivariate data using preattentive processing, *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 5, (1995), 190-221.
- [10] Kay, A.C., Wheeler, S.C., Bargh, J.A., & Ross, L. Material Priming: The influence of mundane physical objects on situational construal and competitive behavioral choice. *Organizational Behavior and Human Decision Processes*, 95, (2004), 83-96.
- [11] Larson, K., Hazlett, R. L., Chaparro, B. S. & Picard, R.W. Measuring the Aesthetics of Reading. In *Proceedings of the British Conference on Human Computer Interactions*. London. (2006).
- [12] Larson, K. The technology of text, *IEEE Spectrum*, 44, 5 (2007), 26-31.
- [13] Lewis, C., & Walker, P. Typographic influences on reading. *British Journal of Psychology*, 80, (1989), 241-257.
- [14] Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & Information Technology*, 25 (2), (2006), 115-126,
- [15] MacLeod, C. The Stroop task in cognitive research. In *Cognitive methods and their application to clinical research*. Eds. A. Wenzel & D. Rubin. APA, Washington, DC, 2005.
- [16] Maison, D., Greenwald, A., & Bruin, R. Predictive Validity of the Implicit Association Test in Studies of

brands, consumer attitudes, and behavior. *Journal of Consumer Psychology*, 14, (2004), 405-415.

- [17] Nisbett, R.E. & Wilson, T.D. Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, (1977), 231-259.
- [18] Palef, S. R. & Olson, D. R. Spatial and verbal rivalry in a Stroop-like task. *Canadian Journal of Psychology*, 29, (1975), 201-209.
- [19] Shaikh, A. D., Chaparro, B. S., & Fox, D. Perception of fonts: Perceived personality traits and uses. *Usability News*, 8(1), (2006).
<http://psychology.wichita.edu/surl/usabilitynews/81/PersonalityofFonts.htm>
- [20] Smith, M. C. & Magee, L. E. Tracing the time course of picture-word processing. *Journal of Experimental Psychology: General*, 109, (1980), 373-392.
- [21] Treisman, A. Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing*, 31, (1985), 156-177.
- [22] Watson, D., Clark, L., & Harkness, A. Structures of Personality and their relevance to psychopathology. *Journal of Abnormal Psychology*, 103, (1994), 18-31.
- [23] Willis, J., & Todorov, A. First Impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17, (2006), 592-597.
- [24] Zajonc, R. Feeling and Thinking: Preferences need no Inferences. *American Psychologist*, 35, (1980), 151-175.