The cover features a light gray background with three large overlapping circles: a white circle at the top, a red circle at the bottom left, and a gray circle at the bottom right. A vertical red line is positioned to the left of the title text.

**KOÇ
UNIVERSITY
UNDERGRADUATE
PSYCHOLOGY
JOURNAL**

ISSUE 2 / 2017

EDITORIAL BOARD

EDITOR-IN-CHIEF:

Eylül Turan

ASSOCIATE EDITORS:

Can Çarkođlu

Erim Kızıldere

Yasin Kađan Porsuk

Esmâ Betül Savaş

Orhun Uluşahin

FACULTY ADVISORS:

Fuat Balcı

Tilbe Göksun

The ideas and opinions contained in this publication solely reflect those of the authors and not the Department of Psychology at KU.

For further information: visit kuupj.ku.edu.tr or email psychologyjournal@ku.edu.tr

TABLE of CONTENTS

ii **Editorial Board**

iv **Editor's Note**

Eylül Turan

EXPERIMENTAL ARTICLES

1 **Emotional Mood and Working Memory: Does Emotion Manipulation Have an Effect on Working Memory Performance?**

Zeynep Adıgüzel, Dila İplik, Naz Rota

13 **How the Visibility of Time Passage Influences Perception of Stress in a Challenging Mental Task**

Eylül Turan, Esma Betül Savaş, Sumru Duraner, Aslı Töre

SCALE DEVELOPMENT ARTICLES

22 **BEEM Emotional Eating Scale: A scale development and validation study**

Banu Bilgili, Edit Barokas, Mert Ege, Ümit Erdem Yiğitoğlu

28 **Development of Premarital Sex Attitude Scale**

Ekin Çakır, Ezgi Yıldız, Dilara Alkan, Meryem Giden, Bengi Türer

REVIEW ARTICLES

39 **An Insidious Stressor: The Concept of Internalized Homonegativity and its Relation to Ethnicity, Religious Settings, Sexual Orientation Disclosure and HIV Risk Behaviors**

Mert Nedim Mercan

48 **The Relation Between Theory of Mind and Language Development: A Research Proposal to Determine the Direction of Causality**

Emre Fatih

EDITOR'S NOTE

It is with great pleasure that I introduce Volume II of the Koç University Undergraduate Psychology Journal. Over the 5-month-period the editors and I had the privilege of exploring some very interesting and insightful articles. Unfortunately we can print only a limited number of them, but we hope to include the remaining submissions in our future issues.

We started this journey with the excitement of forming a platform in which undergraduates can find the opportunity to share their research with fellow students. In this issue we aimed to expand our platform to undergraduates from other universities, and included articles from *Yeditepe University* and *Middle East Technical University*.

I thank the editors for their tireless work on this issue; they have voluntarily reviewed each article sentence by sentence with the goal of helping fellow students to enhance their work and to enrich their research by contributing with new perspectives. I would also like to thank to our advisors Dr. Tilbe Göksun and Dr. Fuat Balcı. Without their support and encouragement, this platform and this issue would not be established. I also want to present my gratitude to the Koç University Psychology Department who supported this journal and this issue throughout the process.

Lastly, I thank to all authors and all undergraduates who shared their research with us, since it is their work that serves as the foundation of this Journal.

As the journal team, we hope that you will find interesting findings, discussions and perspectives in this issue, and more importantly we hope that you will be inspired to share your research with us!

Editor-in-chief

Eylül Turan

Emotional Mood and Working Memory: Does Emotion Manipulation Have an Effect on Working Memory Performance?

Zeynep Adıgüzel, Dila İplik, Naz Rota

Koç University

How does your mood affect your memory? This study aims to investigate the influence of task irrelevant emotions on working memory performance. Our main interest was regarding whether emotional state of the individual had a strengthening or weakening effect on working memory performance. Participants were assigned to one of three conditions: two experimental conditions (positive and negative movie clips) and a control condition (neutral movie clip). Our experiment consisted of 2 sessions implemented within 2 days, mostly consecutive days. In the first session, in order to measure individuals' baseline working memory span, participants who were between ages of 18-24 (n=41) completed the OSPAN task. In the second session, participants were asked to complete the OSPAN task after being induced with either positive negative or neutral mood states via video clips. We found no significant effects of individuals' induced mood states on their working memory performances. Participants' OSPAN scores did not change depending on the induced mood states.

Keywords: emotions, working memory, induced mood, operation span task

There is a wide array of research which focuses on the relationship between memory and emotions. However, there seems to be different explanations about the relationship and its impact on human cognition. It has been shown that various parameters can influence memory performance. However, emotions seem to interact differentially across memory types (Phelps, 2004). For example, the way emotions interact with episodic memory is different than the way they interact with working memory. Studies investigating the relationship between emotion and memory mainly focused on episodic memories (Holland, & Kensinger,

2010). In contrast, there is relatively less research examining the interaction between emotions and working memory. Working memory is responsible for short-term maintenance, rehearsal and manipulation of information (Baddeley, 2003). The relationship between emotions and working memory has been examined through two dimensions (Phelps, 2004). One dimension is valance, which refers to pleasantness or unpleasantness of an emotion, and the other is arousal, which examines the level of excitement or calmness evoked by emotions. Whether a stimulus evokes pleasant or unpleasant emotions and

whether emotion is low or high in terms of arousal level is important because they have different influences on working memory performance. This study aimed to examine the effects of individuals' moods on working memory performance by inducing participants with either positive, negative or neutral moods before taking the OSPAN task and comparing it with their baseline OSPAN performance's.

Working memory measures an individual's capacity for holding and manipulating the information for short periods of time. However, it has a limited capacity (Baddeley, & Hitch, 1994). It is argued that this limitation might be due to the inability to focus attention on more than four items at a given time. According to Kane et al., (2001) focused attention assists working memory both at encoding and retrieving task relevant stimuli. Thus, focused attention is also vital for optimal working memory performance. In fact, Kane and his colleagues found that there is a strong association between individuals' working memory performances and their attention span. Participants with high working memory span scored higher in both the prosaccade, in which a visual cue and a subsequent to-be-identified target letter are presented in the same location, and antisaccade task, in which a target letter is presented in the opposite location of a visual cue's location, compared to participants with low working memory span (Kane, Bleckley, Conway, & Engle, 2001). In another study, Bleckley and his colleagues (2001) found that participants who scored high in the OSPAN differed from participants who scored low with respect to how they allocate their attention. Also, in Egly and Homa's (1984) selective attention task, participants with low working memory capacity responded slower when the target appeared far from the cue, suggesting

that individuals with higher working memory capacity were able to focus on a wider range more successfully (Bleckley et al., 2003).

Adding to the role of attention in limited nature of working memory, behavioral and neuroimaging research shows that an important factor in allocating attention seems to be emotional aspects of information. More specifically, attention prioritizes emotionally salient stimuli. Stimulus with emotional content captures attention more quickly compared to neutral stimuli (Öhman, Lundqvist, & Esteves, 2001). For instance, in a visual detection task, target stimulus was found faster when it included emotional content compared to neutral stimuli (Öhman, Lundqvist, & Esteves, 2001). Furthermore, once attention is captured by emotional stimuli, it takes longer to disengage from it. In a dot probe task, when target stimulus was cued by an emotional stimulus, participants responded faster compared to when target stimulus was cued by a neutral stimulus (Brosch, & Van Bavel, 2012). However, emotion processing could challenge focused attention by directing attention through emotions instead of task relevant stimuli, which in turn can disrupt working memory performance (Jefferies, Smilek, Eich, & Enns, 2008).

Since there are various conflicting results present, the literature lacks consensus on the possible interactions. Differences in findings could be due to a) participants' ages, b) whether emotionally charged material is presented as task relevant or as a distractor and c) whether the evoked emotions are positive or negative. Research findings indicated that both young and old participants' performances improved when target stimulus was emotionally charged compared to neutral target stimulus. However, disruptive effects of negatively

charged distractors were only observed in performance in older adults (Truong & Yang, 2014). Previous neuroimaging studies showed that these distinct contributions of positive and negative emotions to working memory could also be traced through neural networks by measuring brain activity using fMRI during encoding and retrieving phases in the Reading Span Task (RST) with using sentences elicited positive, negative or neutral emotions. Participants were asked to read these emotional sentences while they were also memorizing target words from sentences. When participants' brain activations were compared, significant activation was found in the right amygdala and parahippocampal gyrus in the negative RST compared to the neutral RST during the reading phase. In contrast, in the positive RST, compared to the neutral RST, significant activation in the substantia nigra was detected during the reading phase. (Osaka, Yaoi, Minamoto, & Osaka, 2013). These results indicate that positive and negative emotions interact with working memory and attention through different neural pathways.

To further investigate whether different emotional states yield a varying working memory performance, Choi et al., (2013) studied whether three levels of emotional states had different influences on working memory performance. Participants were presented with either neutral, tense or relaxed emotion evoking pictures between 3-back working memory tasks. Findings illustrated that participants who were in the neutral picture condition scored highest followed by the relaxed and tense conditions which suggests that emotion eliciting pictures might interfere with working memory and disrupt the process (Choi et al., 2013). In contrast, another study examining the effects of emotion processing on working

memory found that negative emotions facilitated interference resolution (Levens & Phelps, 2008). In a modified version of a recent probe paradigm participants' reaction times were compared between interference and non-interference trials for either words with negative content or neutral words. Participants in the interference condition with negative words responded faster than those in the interference condition with neutral words, which points out that emotion processing could play a role in interference resolution. Furthermore, Storback and colleagues (2015) examined the influence of emotion processing on domain specific working memory and they induced positive, negative or neutral mood to participants before they completed a verbal operation span task and a spatial working memory task. Participants who were induced with positive mood before working memory tasks performed better on both domains compared to negative and neutral mood conditions (Storbeck & Maswood, 2015). Interestingly, another study suggested that emotion processing could impair the working memory performance by spreading the activation among items in working memory and making it more diffuse. Therefore, items in the working memory might not be able to reach the threshold activation level, making their retrieval more difficult (Davelaar, GoshenGottstein, Ashkenazi, Haarmann, & Usher, 2005). Martin et al. revealed that when participants' moods were manipulated by presenting them with positive videos before the completion of the Running memory span task, participants in the positive mood induction condition performed worse in the working memory task compared to participants in the neutral mood condition. (Martin, & Kerns, 2011). However, contradictory evidence was presented in a study

which suggested that emotion processing facilitated working memory performance. Using a visual 2-back task participants were assigned to either positive, negative or neutral stimuli conditions. Participants who were exposed to emotionally charged pictures (both negative and positive) performed better than those who were exposed to neutral pictures between trials. In contrast to studies providing evidence for processing emotionally charged stimuli either facilitating or disrupting working memory performance, there were studies which could not find any interactions among them. Erk and colleagues (2007) designed a task in which participants completed a modified Steinberg recognition task. Delay conditions were manipulated in the task and participants were presented with either a blank screen, a positive picture or a negative picture. Their results indicated no significant difference in working memory performances of participants (Erk, Kleczar & Walter, 2007).

As shown, there seems to be contradictory evidence present in the literature. While some researchers claim that emotion processing disrupts working memory performance by manipulation attention through emotions, others suggest that emotion processing enhance working memory performance.

This study examined how emotion processing and induced mood interact with working memory performance. As mentioned above, previous studies mostly investigated the influence of emotionally charged material on working memory performance and revealed inconsistent results from one another. Further, the number of studies which manipulated the task irrelevant mood of the participants is limited. Considering the limited capacity of the working memory (Baddeley, & Hitch, 1994)

and its dependence on focused attention, we expected to observe a decline in individuals' working memory performance, hypothesizing that their focuses will shift towards their emotions rather than the presented material.

We aimed to expand and contribute to the literature by manipulating individuals' moods before conducting a working memory task and investigating whether the emotional state of the individual has a strengthening or weakening effect on working memory performance. We attempted to manipulate the emotional states of the participants by making them watch short movie clips, and (then) measuring their working memory performances both before (baseline) and after emotional manipulation.

Method

Participants

Younger adults with an age range of 17-24 years participated in this study (Mean age=21.66, SD=1.37). Snowball sampling was used to recruit participants. They were consisted of 23 males and 18 females. All of our participants were from Koç University, except from one participant. All of our participants were unmarried. Education levels of participants were divided into three categories, namely high-school graduate, university student and university graduate. Job status was defined as either "employed" or "not employed." Due to the repeated-measured design of this experiment, participants took part in the experiment for 2 consecutive days. The experiment consisted of 2 sessions and each session took approximately 30 minutes. The total duration of the experiment was about an hour. Participants were randomly assigned to one of 2 experimental (positive or negative) and 1 control group condition (neutral). Due to the

design of the experiment which required repeated measurement of the Operational Span task scores for 2 consecutive days, the number of participants targeted was reduced from 75 to 45. In addition, the data collected from 4 participants were excluded due to experimenter errors.

Materials

Movie Clips. The stimuli for manipulation of emotional state were drawn from a large database of emotion-eliciting films developed by Schaefer et al. (2010): (1) high arousal positive: “*Benny and Joon*” (3.80) Benny (Johnny Depp) plays the fool in a coffee shop, (2) high arousal negative: “*A Perfect World*” (4:39) Butch (Kevin Costner) is gunned down, at the end of the movie, and (3) neutral stimulus: “*The Lover*”(0.43) The young lady (Jane March) is walking in a crowded area getting into the car.

Measures

Demographics. Participants were asked to indicate their age, sex, marital status (married, single, divorced, widow), educational level (university student, university graduate and high school student) and job status (whether they have a paid job or not). Participants who had a part-time job were also regarded as the ones who had a paid job.

General Emotion Assessment Scale. Participants’ general emotional states were measured with self-reported General Emotion Assessment Scale (GEAS) created by the authors. Participants were asked to state their general emotional states, not according to how they felt during the clips, but according to how they feel in general, on a 6-item mood scale: happiness, sadness, anger, fear, excitement and disgust. Each mood item was evaluated using a 5-point Likert scale from (1): absolutely not agreed, to (5) Absolutely agreed. With data

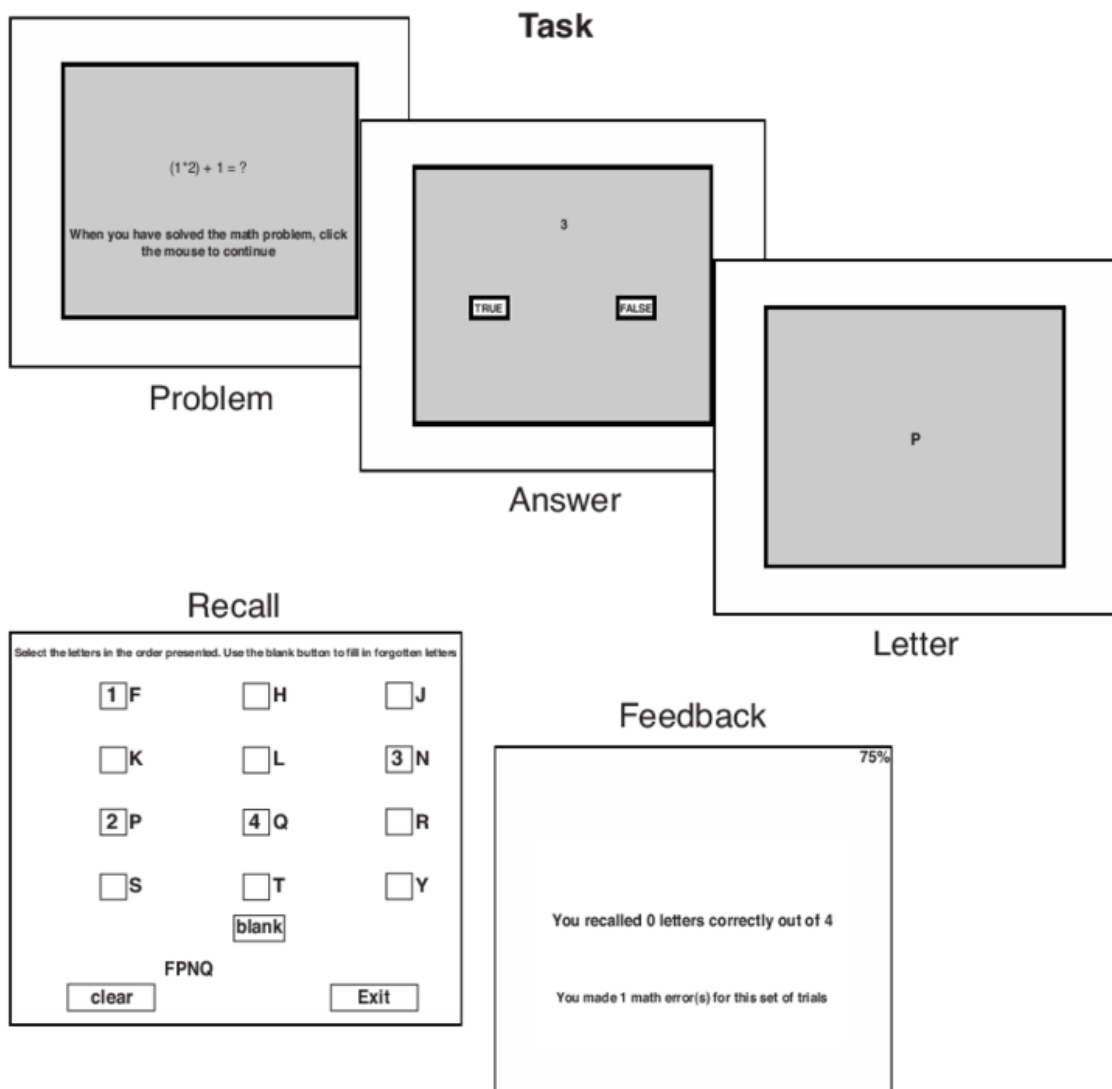
obtained using this self-reported scale, we then created 2 different subscales. Using 2 items (*happy* and *excited*) of the GEA that were related with positive emotions, we created a Positive Emotion Assessment Subscale (PEAS) and 4 items (*anger*, *fear*, *excitement* and *disgust*) of the GEA that were related with negative emotions, we created a Negative Emotion Assessment Subscale (NEAS).

Self- Reported Movie Evaluation Scale. Since we were interested in manipulating the emotional states of participants via presenting them with short movie clips and measuring their working memory performances before and after emotional manipulation, we adapted the “Self-reported emotional arousal scale” of Schaefer et al. (2010) and developed a 2-items scale to measure the intensity of emotional arousal, also as a manipulation check: First item was measuring subjective emotional arousal using a 5-point Likert scale from 1-not all aroused to 5-very aroused: *Filmi izlerken hissettiğiniz heyecan düzeyi nasıldı?* (How would you rate your arousal level while watching the movie?) Second item was measuring the subjective evaluation of the movie clip’s emotional content using as well a 5-point Likert scale from 1-very negative to 5-very positive : *Filmi izlerken nasıl duygular hissettiniz? (What kind of emotions did you feel during watching clips?)* Participants were kindly asked to fill out this scale according to what they actually felt while watching the movie clips, rather than what they thought they should feel.

Operation Span Task (OSPAN). In order to measure working memory performance, we used an automated version of the operation span task developed by Unsworth et al. (2005), which is proven to be reliable and valid (see Figure 1). This task is easy to implement since

it is mouse-driven and records the results itself. At the end of the task, the calculated results were presented on the computer screen. The program operates on the basis of the average time of an individual to solve the math operations plus 2.5 *SD*, which was indicated by Unsworth et al. (2005) that it was enough time to prevent participants from rehearsing and allow them to work at their own unique pace. At the end of each set of math operations, participants were asked to recall the letters by putting them in order. The OSPAN also gives feedback of cumulative accuracy of math operations and number of items recalled in each

set (Unsworth et al., 2005). The accuracy threshold is set at 85% in order to prevent participants from dedicating all of their cognitive resources to recalling the letters. In the end, the program calculates two span scores: absolute span, which denotes the sum of set sizes that are correctly recalled, and total correct, which denotes the total number of letters that are recalled in the correct place. In addition, what the program also reports is the number of total math errors, which can be divided as (1) the number of math errors in which participants spend more time than they are allowed to solve the math operation, and (2)



the number of accuracy errors in which the participant fails to solve the math operation correctly (Unsworth et al., 2005).

Design

The research design of the study was a 2x3 mixed design and it studied the relationship between emotional state of the individual and the working memory performance. The variables of this study consisted of assessment of OSPAN at different time points and emotional manipulation condition. Participants are assigned to one of the 3 conditions: 2 experimental conditions (positive and negative movie clips) and the control condition (neutral movie clip).

Procedure

The experiment consisted of 2 sessions (Session 1 and Session 2). Sessions were implemented within 2 consecutive days. Both sessions were done in a quiet environment in the library of the university. Instructions were standardized by the three experimenters. Confidentiality was assured and the all of the forms and results were only accessible by the experimenters.

Session 1: As the participants arrived, they were asked to carefully read the informed consent form and to sign it if they agreed to participate in the experiment. They were also verbally informed that they could leave the experiment at any phase they wanted. If the participants agreed to continue the experiment, they were asked to fill out the demographic form. The experimenters then presented the participants with the General Emotion Assessment Scale (GEAS) and instructed the participants to answer the questions depending on how they felt in general. After these surveys were completed, the participants were placed in front of a computer to complete the Operation Span Task (OSPAN) task, which took

approximately 20 minutes. Each time, before starting the actual task, there was a practice trial. Participants were informed about the task and asked to call the experimenter when they saw the screen stating to do so. Session 1 was ended when participants completed OSPAN.

Session 2: Participants who took part in Session 1 were invited again in order to complete the second part of the experiment. It is important to mention that Session 2 was conducted within 2 days. Out of our initial sample of 45 people, 91.1 % of participants participated in the following day while 8.9 % of the participants participated within 2 days. After watching a movie clip, participants were asked to complete the Movie Evaluation Scale based on how they felt while watching the movie clip, rather than how they felt in general. Once the evaluation scale was filled out, participants were asked to complete the OSPAN which they were familiar with from the previous session. Session 2 ended when they completed the OSPAN for the second time. Finally, participants were debriefed about the actual goal of the experiment.

Results

Our primary approach was to use scores of our Positive and Negative Emotion Assessment Subscales (PEAS and NEAS) as the covariates, and then to investigate the impact of emotional manipulation condition (Positive, Negative, and Neutral movie clips) across the 2 OSPAN Absolute Scores assessments (with time of assessment as a repeated measure). We conducted a 2x3 ANCOVA with time of OSPAN assessment (Session 1 and Session 2) as a within subject variable and emotional manipulation condition (Positive, Negative and Neutral movie clips) as

a between subject variable. Analysis of covariance showed that there was no significant main effect of time of OSPAN assessment and the emotion manipulation condition, $F(1, 36)=.272, p=.605, \eta^2=.008$, $F(2, 36)=.945, p=.398, \eta^2=.050$, respectively. There was also no significant interaction between time of OSPAN assessment and emotional manipulation condition, $F(2, 36)=1.463, p=.245, \eta^2=.075$. When we also controlled for sex of the participants, total scores of General Emotion Assessment Scale (GEAS) and Movie Evolution Scale Item 1 (measure of valence manipulation) and 2 (measure of arousal manipulation), results did not differ.

However, when we conducted a correlational analysis using the Pearson correlation coefficient between the 2 items of Self- Reported Movie Evaluation Scale: measure of valence manipulation (Movie Evolution Scale Item 1) and measure of arousal manipulation (Movie Evolution Scale Item 2),

we found a significant positive correlation between these 2 items ($r=.321, p=.041$). Also, was negatively correlated with scores of Positive Emotion Assessment Subscale (PEAS), but positively correlated with scores of Negative Emotion Assessment Subscale (NEAS), $r=-.338, p=.03$ and $r=.314, p=.046$, respectively. Measure of valence manipulation (Item 2) was marginally significantly correlated with scores of Positive Emotion Assessment Subscale (PEAS), ($r=-.262, p=.098$), but not correlated with scores of Negative Emotion Assessment Subscale (NEAS), ($r=.217, p=.174$). Further analysis showed that, for neutral emotion manipulation condition, there was a significant negative correlation between scores of Positive Emotion Assessment Subscale (PEAS) and measure of arousal manipulation (Item 1), ($r=-.521, p=.047$), and a marginally significant negative correlation between scores of Positive Emotion Assessment Subscale (PEAS) and measure of valence manipulation

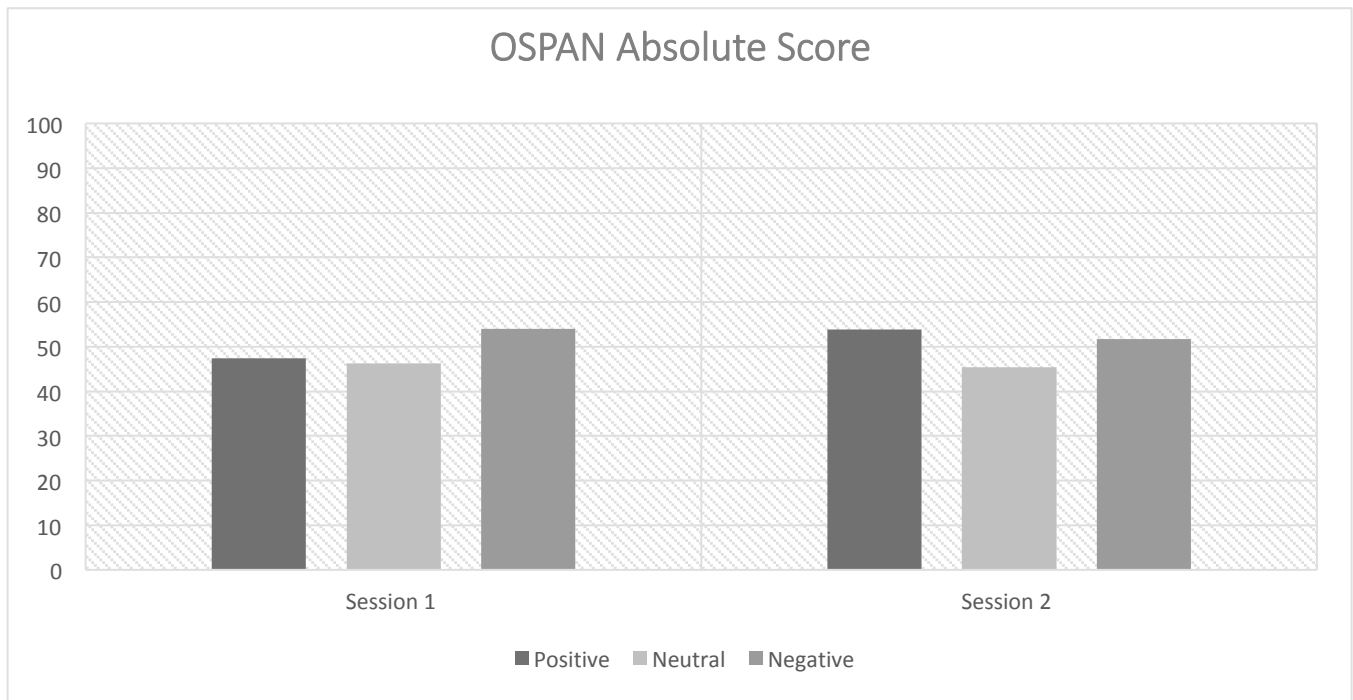


Figure 2: Mean of participant's OSPAN absolute scores from Session 1 to Session 2. In Session 1, participants aren't exposed to the positive, neutral and negative movie clips yet. Since participants in each emotional manipulation group in Session 1 are going to watch it in Session 2, experiment and control groups are named as positive, neutral and negative group for Session 1 either

(Item 2), ($r = -.512, p = .051$).

One-way ANOVA analysis showed that participants who watched one of the movie clips (positive, negative and neutral) did not differ in their OSPAN Session 1 and Session 2 Absolute Scores, $F(2, 38) = .983, p = .383$ and $F(2, 38) = 1.044, p = .362$, respectively. Results did not change for both OSPAN Total Session 1 and Session 2, and Math Errors Session 1 and Session 2 as seen in Figure 2.

Discussion

In this research, we aimed to examine whether emotion processing and induced mood interact with working memory performance. Therefore, we attempted to manipulate participants' moods before asking them to perform the working memory task. After mood manipulation, they were given a working memory task again. Our main question was whether the emotional state of individuals has a strengthening or weakening effect on their working memory performances.

In this study, we found no relationship between induced moods and their effects on working memory performance. Our results showed that there is not enough evidence to suggest that processing an emotionally arousing stimulus strengthens or weakens working memory performance. There might be several explanations that can be discussed in the light of the literature on attention and emotion.

We had hypothesized that participants' working memory performances would be higher when they were presented with emotional stimuli, e.g., positive or negative movie clips, compared to neutral stimuli. There is accumulated evidence showing that working memory is a crucial component for several cognitive processes such as complex decision making and selective attention operations

(Baddeley, 1986). By looking at the strong relationship between working memory performance and attention span, it can be suggested that working memory is affected by attention-based properties. Also, emotions are suggested to have a role in directing and grabbing attention (Oatley & Jenkins, 1996). However, our study provides results conflicting with such findings since emotional manipulation on individuals did not yield any difference on their working memory performance in the present study. As Jeffrey et al. (2008) suggested, we could argue that emotion processing shifted focused attention from task relevant stimuli to emotional stimuli, which, in turn, might have impaired the working memory performance of our participants.

The suggestion that emotional processing has a disruptive effect on working memory can also be explained by previous findings indicating that regions of the prefrontal cortex are related with emotional working memory tasks through different pathways (Grimm et al., 2012). In their research, Grimm et al. (2012) provided evidence for emotional processing leading to an increase in the activation of cognition-related lateral prefrontal regions, while cognitive effort lead to decrease in activation in emotion-related cortical midline regions.

It is also important to mention our ability to manipulate emotions of our participants. We used 3 movie clips (positive, negative and neutral) and participants who were in one of the 3 emotional manipulation conditions did not display significant differences between their session 1 and session 2 OSPAN performances. This could be due to our inability to manipulate their emotions to the extent that we had hoped for. The movie clips

that we used might have been insufficient in terms of evoking emotions or participants simply might have been preoccupied with anything in their lives. Future research could benefit from using longer and more arousing movie clips for a more effective mood manipulation.

There are two popular dimensions in examining emotions and how they relate to working memory (Phelps, 2004). Since we found a significant positive correlation between the measure of valence manipulation (Movie Evolution Scale Item 1) and the measure of arousal manipulation (Movie Evolution Scale Item 2), it can be argued that the way we approach emotions and their relationship with working memory is in line with the literature. The more a stimulus evoked pleasant emotions, the higher it was in arousal. This finding should support the way the literature evolves around emotion and working memory research.

However, it also important to note that measure of arousal manipulation (Item 1) was negatively correlated with scores of Positive Emotion Assessment Subscale (PEAS), but positively correlated with scores of Negative Emotion Assessment Subscale (NEAS). One possibility might be that arousal manipulation had a negative connotation more than a positive one. Perhaps our negative emotion manipulation clip was at higher levels of arousal than positive emotion manipulation clip. Although we thought that it was unlikely to explain this finding, future research should consider how arousal is particularly related with positive and negative emotions

The finding that, for the neutral emotion manipulation condition, scores of Positive Emotion Assessment Subscale (PEAS) were negatively correlated with both the measure of arousal manipulation (Item 1) and the measure

of valence manipulation (Item 2) could be interesting to discuss. For participants who were in the control condition, the higher they scored on the Positive Emotion Assessment Subscale (PEAS), the less they scored on valence and arousal measurements with the neutral movie clip. Even though it may seem like arousal and valence have a negative connotation, it would not be sufficient to claim as such.

Limitations

There are at least four known limitations of the current study. First, since time of assessment was a within variable, we had to take the risk of facing a practice effect (i.e., Getting used to the task when performing it for the second time). Another issue with time of assessment of working memory task was the fact that two sessions of the task were not collected at the same time of each day. Second limitation is our sample. Since participants are mainly undergraduate students at Koç University, the sample might not be representative of the adult population. We collected data from the most convenience sample. Also sample size was small. Another limitation is that the duration of movie clips that we used for emotional manipulation, which failed to induce mood. We may have had better results if we would have used longer movie clips with a shorter memory task. Finally, we gained people's general emotion states that relied on self-report, which can increase the social desirability bias. People may have answered the questions in a way that they thought they should.

In conclusion, we hypothesized that when participants were presented with emotional manipulation (movie clips), their working memory performances would be higher when they were presented with emotional stimuli, e.g., positive or negative

movie clips, compared to neutral stimuli. However, our study provided evidence for how emotional manipulation of participants did not result in any difference on their working memory performances. We could not find any meaningful relationship between inducing mood on participants and their working memory performances. Our study provides results conflicting with such findings since emotional manipulation on individuals did not yield any difference on their working memory performances.

References

- Baddeley, A., & Hitch, G. (1994). Developments in the concept of working memory. *Neuropsychology*, 8(4), 485-493. <http://dx.doi.org/10.1037//0894-4105.8.4.485>
- Baddeley, A.D., 1986. Working Memory. Clarendon Press, Oxford University Press, Oxford Oxfordshire, New York.
- Bleckley, M. K., Durso, F. T., Crutchfield, J. M., Engle, R. W., & Khanna, M. M. (2003). Individual differences in working memory capacity predict visual attention allocation. *Psychonomic bulletin & review*, 10(4), 884-889.
- Brosch, T., & Van Bavel, J. (2012). The flexibility of emotional attention: Accessible social identities guide rapid attentional orienting. *Cognition*, 125(2), 309-316. <http://dx.doi.org/10.1016/j.cognition.2012.07.007>
- Choi, M., Min, Y., Kim, H., Kim, J., Yeon, H., & Choi, J. et al. (2013). Effects of three levels of arousal on 3-back working memory task performance. *Cognitive Neuroscience*, 4(1), 1-6. <http://dx.doi.org/10.1080/17588928.2011.634064>
- Davelaar, E., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H., & Usher, M. (2005). The demise of short-term memory revisited: empirical and computational investigations of recency effects. *Psychological Review*, 112(1), 3-42. <http://dx.doi.org/10.1037/0033-295x.112.1.3>
- Erk, S., Kleczar, A., & Walter, H. (2007). Valence-specific regulation effects in a working memory task with emotional context. *Neuroimage*, 37(2), 623-632. <http://dx.doi.org/10.1016/j.neuroimage.2007.05.006>
- Grimm, S., Weigand, A., Kazzer, P., Jacobs, A. M., & Bajbouj, M. (2012). Neural mechanisms underlying the integration of emotion and working memory. *Neuroimage*, 61(4), 1188-1194.
- Holland, A., & Kensinger, E. (2010). Emotion and autobiographical memory. *Physics Of Life Reviews*, 7(1), 88-131. <http://dx.doi.org/10.1016/j.plrev.2010.1.006>
- Jefferies, L., Smilek, D., Eich, E., & Enns, J. (2008). Emotional Valence and Arousal Interact in Attentional Control. *Psychological Science*, 19(3), 290-295. <http://dx.doi.org/10.1111/j.1467-9280.2008.02082.x>

- Kane, M., Bleckley, M., Conway, A., & Engle, R. (2001). A controlled-attention view of working-memory capacity. *Journal Of Experimental Psychology: General*, *130*(2), 169-183. <http://dx.doi.org/10.1037//0096-3445.130.2.169>
- Kensinger, E., & Corkin, S. (2003). Effect of negative emotional content on working memory and long-term memory. *Emotion*, *3*(4), 378-393. <http://dx.doi.org/10.1037/1528-3542.3.4.378>
- Levens, S., & Phelps, E. (2008). Emotion processing effects on interference resolution in working memory. *Emotion*, *8*(2), 267-280. <http://dx.doi.org/10.1037/1528-3542.8.2.267>
- Lindström, B., & Bohlin, G. (2011). Emotion processing facilitates working memory performance. *Cognition & Emotion*, *25*(7), 1196-1204. <http://dx.doi.org/10.1080/02699931.2010.527703>
- Martin, E., & Kerns, J. (2011). The influence of positive mood on different aspects of cognitive control. *Cognition & Emotion*, *25*(2), 265-279. <http://dx.doi.org/10.1080/02699931.2010.491652>
- Oatley, K., & Jenkins, J.M. (1996). Understanding emotion. Cambridge, MA: Blackwell Publishers.
- Osaka, M., Yaoi, K., Minamoto, T., & Osaka, N. (2013). When do negative and positive emotions modulate working memory performance?. *Scientific Reports*, *3*(1). <http://dx.doi.org/10.1038/srep01375>
- Phelps, E. (2004). Human emotion and memory: interactions of the amygdala and hippocampal complex. *Current Opinion In Neurobiology*, *14*(2), 198-202. <http://dx.doi.org/10.1016/j.conb.2004.03.015>
- Storbeck, J., & Maswood, R. (2015). Happiness increases verbal and spatial working memory capacity where sadness does not: Emotion, working memory and executive control. *Cognition And Emotion*, *30*(5), 925-938. <http://dx.doi.org/10.1080/02699931.2015.1034091>
- Truong, L., & Yang, L. (2014). Friend or foe? Decoding the facilitative and disruptive effects of emotion on working memory in younger and older adults. *Frontiers In Psychology*, *5*. <http://dx.doi.org/10.3389/fpsyg.2014.00094>
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior research methods*, *37*(3), 498-505.

How the Visibility of Time Passage Influences Perception of Stress in a Challenging Mental Task

Eylül Turan, Esma Betül Savaş, Sumru Duraner, Aslı Töre

Koç University

The effect of time constraint on stress and performance in a mental rotation task was investigated in an experimental setting with no time constraint, time constraint, and time constraint with a visible countdown clock conditions. Participants were 73 university students from Turkey. The findings indicated that there is a significant increase in stress levels of the participants between no time constraint and time constraint with a visible countdown clock conditions. This demonstrated that tracking a clock while solving a challenging mental task caused distress in the participants. Performance, on the other hand, decreased across the three conditions; participants performed best in the no time constraint condition and worst in the time constraint with a visible clock condition. Thus, further studies can investigate the mechanisms behind time constraint and its possible indications on stress and performance.

Keywords: time constraint, psychological stress, performance, cognitive task

Psychological stress has been a popular subject of research due to its implications on daily life and functioning of the individuals. It has been shown that chronic stress can have immense effects on cognitive functioning and mental health (Marin et al., 2011). Therefore, the importance given to studying psychological stress increases day by day with the hope to understand its mechanisms. There are physiological indications of stress such as stress hormones, e.g. glucocorticoids and catecholamine, and these are shown to have an effect on cognitive functions (Lupien, 2007). Psychological stress can show itself in many forms, and time stress can be considered one of the most popular of its subcategories. This article will especially focus on stress induced by time.

One feels stressed when time given does not seem to be enough to perform the tasks they want or need to perform. Time constraint triggers physiological responses mentioned before, which result in the feelings of stress. The body's response to stress may vary across individuals (Lupien, 2007). This is aligned with the fact that people have different levels of stress proneness in general, as well as different reactions to it. In her review article, Wilson (2002) investigates the claim that cognition is time pressured and states that time pressure is thought to create a representational bottleneck, which constrains humans' cognitive capacities in the face of a demanding task. Therefore, this could be tied to the relation between time constraint and performance.

Wilke (1985) argues that stress and

productivity are related to each other with an inverted U-function, meaning that productivity is poorest at the lowest and highest levels of stress. However, most of the existing studies regarding the effect of time pressure on performance indicate that time constraint results in a significant decrease in performance. These studies differ in the method they used to measure performance. While some studies use real life situations such as university exams (De Paola & Gioia, 2016), others use cognitive and perceptual tasks (Szalma et al., 2008) and measured accuracy or reaction time as indicators of performance. In all of the relevant studies, participants were assigned to either limited or unlimited time conditions to investigate how time constraint affects performance in these tasks. It is claimed that regardless of the type of the task given to participants, a decrease in performance is observed (Szalma et al., 2008). This decrease differs among participants with different levels of experience on the given task (Ahituv et al., 1998). The effect of time constraint on people's expectations about their performances was investigated in the literature as well. As in the case of performance, time pressure caused lower expectations of their own outcome. A sex difference was also observed, where the negative effect of time pressure was greater in females (De Paola & Gioia, 2016).

Time, also as a dimension in which people make decisions every day, influences perceived stress in situations with a variety of options, and therefore the decision-making strategies of the individuals. It is found that time limit, manipulated in conjunction with the number of alternatives such as not having a time limit, leads to greater levels of reported frustration and perceived decision-related difficulty of the task (Haynes, 2009). How time pressure is a determinant of the risk-

taking behavior, and the strategy used to cope with the risky condition has also been a subject of investigation. It is concluded that individuals are less inclined to take risks, and focus more on the negative consequences under higher time pressure (Zur & Shlomo, 1981). Interestingly, a warning at the beginning of a decision-making task about the insufficiency of time to complete the task is enough to lower the performance of the participants throughout the assignment (DeDonno, 2008). Considering that time-stress causes information-processing overload which in turn results in a decrease in the optimality of decision processes (Arieli, 2001), we expect time pressure to have a negative effect on the performance of our participants in the Mental Rotation Task, developed by Shepard and Metzler (1971). By choosing this test, we aimed to eliminate the potential practice and learning effects for such effects would constitute a confounding variable in our research, and to present a task with moderate challenge to our participants.

While there is a multitude of findings in literature on the effect of time constraint on various concepts such as creativity (Amabile et al., 2002), risk-taking (Zur & Breznitz, 1981; Huber & Kunz, 2007), decision-making (Ariely & Zakay, 2001), and stress (Wilson, 2002), these studies are usually conducted with two conditions: time restriction and no time restriction. And, whether the visibility of the clock affects the participants' responses in the condition of time restriction has received little attention. This addition to the literature stems from the idea that the 'visibility of the clock' condition is something that we encounter in our lives, especially in educational context. Most of the exams take place in an environment in which examinees can track the remaining time from a clock, however it is not thoroughly investigated how this tracking of the clock affects examinees'

performance and stress levels. Therefore, in this research, we aim to investigate whether the visibility of the clock and the constant time tracking has an effect on the stress levels of the participants compared to the task completed with no time restriction or time restriction without a visible clock. We hypothesize that the stress levels of the participants would increase with time constraint and the most stress will be experienced when there is a visible clock that can be tracked, since the clock serves as a constant reminder of the time passing, whereas reminding the remaining time verbally takes place less frequently. We also hypothesize that the performance levels of the participants will change between three conditions.

Methods

Participants

Participants were 73 undergraduate students ($M_{\text{age}}=21.19$, $SD=1.552$, 60% female) from Koç University, who voluntarily participated in the experiment in exchange for 0.5 extra credit points for one of their psychology classes. Assignment to conditions was random. All participants first received the baseline condition that had no time restriction, afterwards they were assigned to one of the three conditions: 25 of them received no time restriction, 24 of them received time restriction without a clock, and 24 of them received time restriction with a visible countdown clock located in front of them. No discard was made among participants.

Measures

General Inventory. Eight questions probing competitiveness orientation, and demographics, and sixteen questions -not reverse coded- probing stress proneness that were chosen from a larger questionnaire

(Eskin et al., 2013). In the questions chosen from the larger questionnaire, the participants rated from 0 (least applicable) to 3 (most applicable) for how related they felt to the situation given in the question (see Appendix A).

Mental Rotation Task. Two sets (baseline and experiment) included nine pairs of images that were either the correct rotated versions of each other or not. The figures were taken from the website “FigShare” (Ganis & Kievit, 2016).

Post-Task Questionnaire. Twelve items including questions probing self-reported stress experienced during the task, plus their preferences of time constraint methods in daily life (see Appendix B).

Time Constraint Conditions

Baseline. Participants were asked to finish the task and given no time constraint in the baseline condition in order to measure their standard accuracy in the test as a base to compare with the experimental condition. All of the participants were given this condition.

No Time Constraint. Identical to the baseline condition, participants were asked to finish the task and given no time constraint in the first condition.

Time Constraint without a visible clock. Participants were told they had 1 minute to finish the task in the second condition. They were notified when the countdown as over.

Time Constraint with a visible countdown clock. A digital clock was used to show the 1-minute constraint by counting down while the participants were engaged with the task in the third condition. The clock rang when the minute was over.

Procedure

A pilot study was conducted with ten participants to test 1) how long it took them to solve the mental rotation tasks in order for us to put a time restriction accordingly, 2) whether there was a difference in difficulty

between baseline and experimental mental rotation task sets, 3) to get feedback on the instruction sheet given. These participants were only given the mental rotation task, and were not asked to fill out the questionnaires. The participants completed the task approximately in 2 minutes, which led us to put a 1-minute constraint in the experiment, making the task difficult but not impossible to finish. The sets were told to be in equal difficulty. And the instruction sheet was told to be understandable.

Participants were tested individually between 10 AM to 2.30 PM to control for the effect of daytime on stress. Upon arriving at the experimentation room, which was selected from the silent rooms in the library of the university, participants were assigned to one of three conditions using simple randomization. The informed consent form was presented to all of the participants. The session started with the "General Inventory" after they signed the informed consent form. After the completion of the inventory, all of the participants were given the instruction sheet of the mental rotation task. This was done to decrease the effect of experimenter wording differences as well as to minimize the interaction between the experimenter and the participant which would minimize the difference in attitudes. After getting approval from the participants that they understood the instructions, they were presented with the baseline condition in which they solved a set of 9 mental rotation questions with increasing difficulty. The order of trials was not counterbalanced. Participants were told that they had no time restriction. After the completion of the first set, they were presented with the second mental rotation task, which included another set of 9 questions with increasing difficulty, similar to the first set. The experimenter informed participants that they would have a) all the

time that they need to finish the task (no time constraint condition) b) 1 minute to finish the task (time constraint without a visible clock condition) c) 1 minute to finish the task and that they can track the remaining time from the clock located in front of them (time constraint with a visible countdown clock condition). The experimenter kept track of time for all conditions while participants were solving the task and informed the participants in the second condition when their time was up. The alarm went off when 1-minute was over in the third condition. After participants finished solving the task, they were asked to complete the post-task questionnaire, which included several questions about their perceived stress levels during and following the completion of the task.

Results

For preliminary analysis, a one-way ANOVA was conducted to test the effect of sex on performance in the Mental Rotation Task. Male participants ($M=91.57$, $SD=10.56$) scored significantly higher than female participants ($M=83.08$, $SD=19.83$) in the baseline set, $F(1,71)=4.465$, $p=.038$, $\eta^2=.059$. However, there was no significant difference between males ($M=82.37$, $SD=18.18$) and females ($M=75$, $SD=21.18$) in the experimental set, $F(1,71)=2.364$, $p=.129$, $\eta^2=.032$, nor in their total performance level –baseline minus experimental set-, $F(1,71)=.051$, $p=.822$, $\eta^2=.001$.

Another additional one-way ANOVA showed that there was a significant difference between two sexes in their stress proneness, such that females' stress proneness ($M=1.47$, $SD=0.531$) was higher than males' ($M=1.20$, $SD=0.527$), $F(1,72)=4.456$, $p=.038$, $\eta^2=.059$. Also, there was no significant effect of faculty on neither performance, $F(5,67)=.388$,

$p=.856$, $\eta^2=.028$, nor perceived stress level, $F(5,67)=.410$, $p=.840$, $\eta^2=.030$.

First independent variable was the time condition with three levels (no time constraint –unlimited time to solve the task-, time constraint without a visible clock –one minute constraint to solve the task-, time constraint with a visible countdown clock – one minute constraint which can be tracked from the clock-). Stress proneness was the second independent variable with two levels (low stress level and high stress level).

A 3x2 between-subjects ANOVA was conducted between three conditions and the stress proneness of the participants. The dependent variable was the change in the stress level of participants between the baseline and experiment sets. There was no main effect of stress proneness, indicating that participants' stress levels as a result of the experiment were not affected by their general stress proneness, $F(1,67)=.079$, $p=.78$, $\eta^2=.001$. However, there was a main effect of condition, as seen in figure 2, $F(2,67)=5.548$, $p=.006$, $\eta^2=.142$. Post-hoc comparisons using Bonferonni correction revealed that there was a significant

difference between the first ($M=2.16$, $SD=0.8$), and the third condition ($M=3.04$, third condition reported significantly higher stress levels compared to the first condition. No interaction was found between the stress proneness and condition type, $F(2,67)=.655$, $p=.523$, $\eta^2=.019$.

To test our second hypothesis, a one-way ANOVA was conducted, which revealed that there was a significant difference among conditions in terms of the performance levels of the participants on the Mental Rotation Task, as seen in figure 3, $F(2,70)=5.40$, $p=0.007$, $\eta^2=.134$. Performance levels were calculated as the percentage of the subtraction of the number of correct answers in the experimental set from the baseline set. Negative outcomes indicated that performance was higher in the baseline set compared to the experimental set. Pairwise comparisons showed that performance in the first condition ($M=-1.77$ $SD=20.20$) significantly differed from both the second ($M=13.42$, $SD=18.23$), and the third ($M=14.35$, $SD=19.51$) condition, $p=0.023$, and $p=0.014$, respectively.

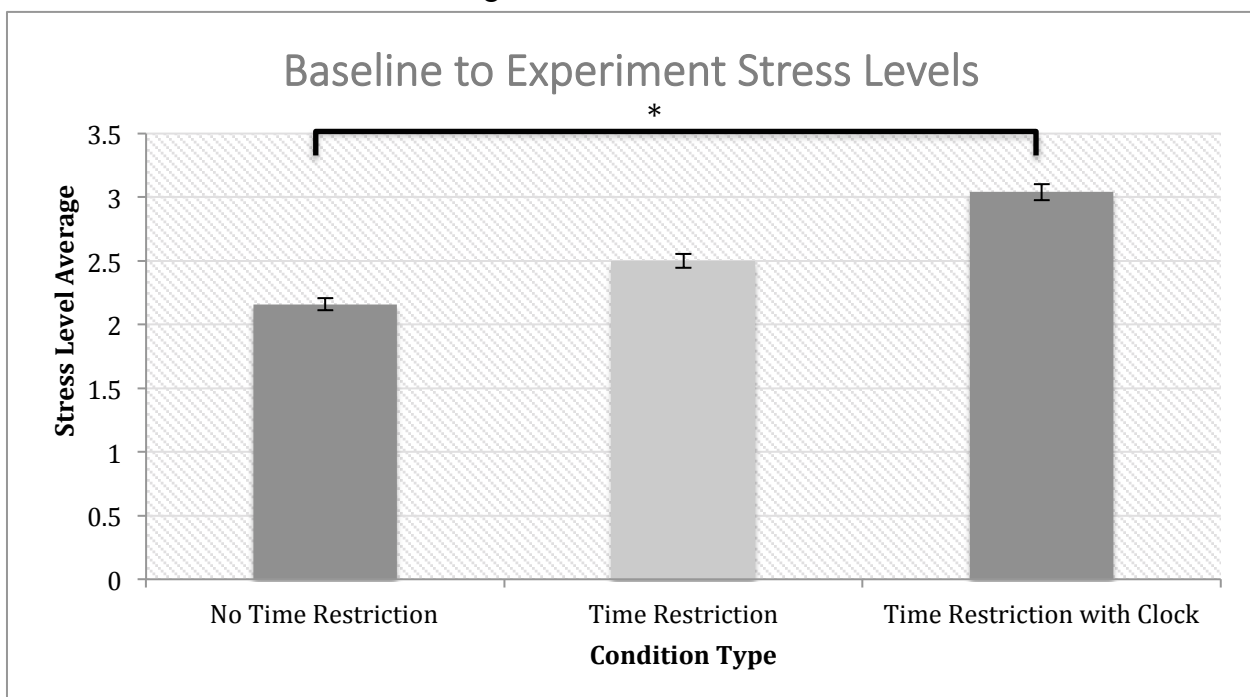


Figure 1. Mean of the participants' baseline to experiment stress levels. It was calculated from a question given in the post-task questionnaire that asked the participants rate their stress levels from 1 to 5, 5 being the highest level.

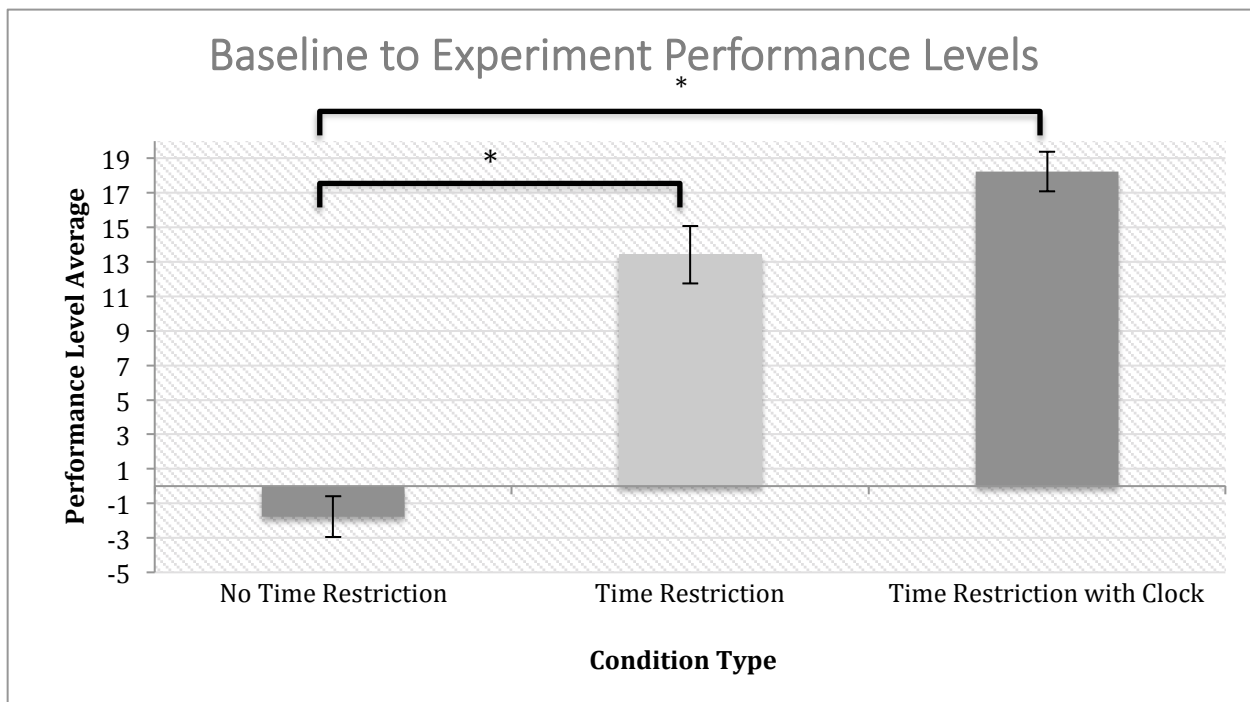


Figure 2. Mean percentages of the participants' baseline to experiment performance levels. It was calculated by extracting the percentage of the correct answers in the experiment set from the correct answers in the baseline set.

A final 2x2 between-subjects ANOVA revealed that even though there was no main effect of sex or stress proneness on perceived stress levels of participants, $p > .05$, there was a cross-over interaction between the two, $F(1,69) = 5.408$, $p = .023$, $\eta^2 = .073$. While females with low levels of stress proneness ($M = 12.50$, $SD = 20.23$) performed better than the high stress proneness ones ($M = 5.55$, $SD = 22.32$), the contrary was found for males. Males with high levels of stress proneness ($M = 17.94$, $SD = 23.36$) performed better than those with low levels ($M = 2.08$, $SD = 10.90$).

Discussion

There is a consensus that time constraint affects performance negatively and puts more pressure on participants by increasing their stress levels (De Paola & Gioia, 2016). Most of the studies conducted on the subject investigate performance and stress levels by giving participants either limited or unlimited amount of time to solve a

task. There are instances that draw an inverted U-function to demonstrate the relationship between stress and productivity (Wilke, 1985). In other words, productivity is poorest at the lowest and highest levels of stress. Although there are instances of inverted U-function, in most cases, performance decreases while the stress level increases when the participants are under time constraint. However, these studies created time constraint using verbal instructions. Participants were told they had x -time to finish the task, and they were informed when their time was over. We hypothesized that time constraint will affect participants' performance and stress levels based on its presentation type, such that the visibility of a clock will increase participants' stress resulting in a decrease in their performance compared to the condition, in which they receive time pressure only by verbal instruction. Our results revealed that participants' performance significantly differed between the first and second, as well

as the first and third conditions, but contrary to what we expected it did not differ between the second and third conditions. We argue that this result could be tied to our sample's characteristics. The fact that college students are accustomed to taking exams under limited durations makes them experienced to perform under time constraint. Whether tracking time on a visible clock or merely being informed about how much time they have to finish a task does not make a difference in their performance. This could be due to their advanced ability to track time even in the absence of a clock.

Previous research suggests that males outperform females on the mental rotation task and have better spatial abilities in general (Baenninger & Newcombe, 1989; Linn & Peterson, 1986; Moore & Johnson, 2008; Parsons et al, 2003). Our results supported this finding only with the baseline set, in which participants did not have any time constraint. However, under time pressure, males and females performed equally, indicating that there is no difference in performance between sexes. This could either mean that females can eliminate the effect of time pressure better than males and close the gap that exists in the no time constraint condition, or it could mean that males have a general tendency to exhibit weakened performance under time constraint.

A counter-intuitive finding in our study, which was the interaction between sex and stress proneness, is an important point of discussion. The fact that males with high stress proneness performed better than the ones with low stress proneness can be interpreted as an indicator that stress becomes an advantage for males, which is not the case for females. We argue that males might be using stress as a motivator to perform better in the task. Due to lack of research on the interaction of sex differences, motivation, and

stress, we think that future studies should investigate these subjects and their possible interactions more thoroughly.

We also investigated whether there is a practice effect by conducting analysis on our control group. There was an increase in their performance from the first set to the second, however this increase was not significant. This finding eliminates alternative explanations such as learning, boredom and response bias.

The major limitation of our study was that our incentive (0.5 extra point for participation) had the potential of being ineffective as they received this extra point regardless of their performance in the experiment. We argue that this is the reason why we could not find a difference between the second and the third conditions. However, as there is a trend seen between the three conditions in their increase of stress and decrease of performance, we argue that there is potential of having a significant difference between time constraint without and time constraint with a clock condition. Future research should either be conducted with the manipulation by providing a more challenging task, or a more effective incentive. Almost all of our participants answered all questions correctly in the baseline set, and participants who had no time restriction got full points in the experimental set as well. Asking them to determine the rotation angle, or making them solve a more complicated Mental Rotation Task could make the task more challenging for them. Also, presenting participants with a monetary reward for their performance would probably increase their motivation level, making the task more demanding.

Another limitation for our study was our inability to record physiological stress responses such as heartbeat or sweating. We think that some people might not have

reported how stressed they actually felt, therefore recording these responses physiologically would be a more objective measure. Future research should include the difference between time constraint with and without a clock conditions while recording the participant's physiological stress responses.

In conclusion, it was our aim to investigate the psychological stress induced by tracking a clock during a task with time constraint. Our results have shown that time constraint is a stressful situation for people when they need to finish a challenging mental task; limited time both increases the stress level, and lowers the performance. Whether individuals visually track the time passage from a clock or are simply informed when time ends is shown not to have a significant difference in terms of their performance. This research paves the way for further studies to examine the effect of time pressure on human cognitive productivity, which might have educational and industrial implications.

References

- Ahituv, N., Igbaria, M., & Sella, A. V. (1998). The effects of time pressure and completeness of information on decision making. *Journal of Management Information Systems*, *15*(2), 153-172. doi: 10.1080/07421222.1998.11518212
- Amabile, T. M., Mueller, J. S., Simpson, W. B., Hadley, C. N., Kramer, S. J., & Fleming, L. (2002). Time pressure and creativity in organizations: A longitudinal field study. *Harvard Business School Working Paper Series*, 02-073.
- Ben Zur, H., Shlomo J. B. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica*, *47*(2), 89-104.
- DeDonno, M. A., & Demaree, H. A. (2008). Perceived time pressure and the Iowa Gambling Task. *Judgment and Decision Making*, *3*(8), 636-640.
- De Paola, M., & Gioia, F. (2016). Who performs better under time pressure? Results from a field experiment. *Journal of Economic Psychology*, *53*, 37-53. doi: 10.1016/j.joep.2015.12.002
- Eskin, M., Harlak, H., Demirkıran, F., & Dereboy, Ç. (2013). Algılanan stres ölçeğinin Türkçe'ye uyarlanması: Güvenirlik ve geçerlik analizi [The adaptation of the perceived stress scale into Turkish: A reliability and validity analysis]. *New/Yeni Symposium Journal*, *51*(3), 132-140.
- Haynes, G. A. (2009). Testing the boundaries of the choice overload phenomenon: The effect of number of options and time pressure on decision difficulty and satisfaction. *Psychology and Marketing*, *26*(3), 204-212. doi: 10.1002/mar
- Kievit, R., & Ganis, G. (2014). *Example stimulus*, Retrieved from https://figshare.com/articles/Example_stimulus/1044175/1
- Lupien, S., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and Cognition*, *65*(3), 209-237. doi:10.1016/j.bandc.2007.02.007
- Linn, M. C., & Petersen, A. C. (1986). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child development*, *56*, 1479-1498.
- Moore, D. S., & Johnson, S. P. (2008).

- Mental rotation in human infants: A sex difference. *Psychological Science*, *19(11)*, 1063-1066. doi: [10.1111/j.1467-9280.2008.02200.x](https://doi.org/10.1111/j.1467-9280.2008.02200.x)
- Marin, M., Lord, C., Andrews, J., Juster, R., Sindi, S., Arsénault-Lapierre, G., Fiocco, A. J., & Lupien, S. J. (2011). Chronic stress, cognitive functioning and mental health. *Neurobiology of Learning and Memory*, *96(4)*, 583-595. doi:10.1016/j.nlm.2011.02.016
- Parsons, T. D., Larson, P., Kratz, K., Thiebaut, M., Bluestein, B., Buckwalter, J. G., & Rizzo, A. A. (2003). Sex differences in mental rotation and spatial rotation in a virtual environment. *Neuropsychologia*, *42(4)*, 555-562. doi:10.1016/j.neuropsychologia.2003.08.014
- Szalma, J. A., Hancock, P.A., & Quinn, S. (2008). A meta-analysis of the effect of time pressure on human performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *52(19)*, 1513-1516. doi: 10.1177/154193120805201944
- Wilson, Margaret. (2002). Six Views Of Embodied Cognition. *Psychonomic Bulletin & Review*, *9(4)*, 625-636. doi:10.3758/BF03196322

BEEM Emotional Eating Scale: A scale development and validation study

Banu Bilgili, Edit Barokas, Mert Ege, Ümit Erdem Yiğitoğlu

Koç University

For humans, eating is not only an action for survival, but also a mean for maintaining a positive mental state. In this study we introduce a scale to measure the degree of psychological significance of eating for individuals. The scale, named BEEM, was formulated on three proposed aspects of eating: physiological hunger, social influences and eating as a mean for coping with psychological stress. Our sample consisted of 162 Turkish participants with ages ranging from 18 to 65. The statistical analyses performed suggest that emotional eating for social compliance, coping, and physiological needs are important dimensions of emotional eating. Furthermore, we provide reliability values for our own items and validity values procured from established scales relating to our variables of interest such as Perceived Stress Scale, Mindful Eating Questionnaire and Marlowe-Crowne Social Desirability Scale. Our analyses suggest that BEEM scale is a valid and reliable scale to measure emotional eating.

Keywords: emotional eating, scale development, validation

Eating is one of the primary physiological needs for many organisms and its satisfaction is vital for survival, growth and maintenance of a properly functioning body. Even though there is an impression of eating that it only appeases hunger, this might not always be the case. Previous research suggests that eating behavior can also be initiated by psychological motivations such as overcoming anxiety, handling distress, and self-indulgence (Jackson, Lynne Cooper, Mintz, & Albino, 2003).

Normally, the bodily response to negative arousal states is an imitation of feeding-induced satiety (Heatherton, Peter Herman & Polivy, 1991). Thereby, when most people face with anxiety or distress, they experience gastric discomfort, which results in suppressed satiety. However, some

individuals respond to emotional agitation in an exact opposite manner, with an enhanced appetite. The psychosomatic view (Kaplan & Kaplan, 1957) describes this type of eating behavior with its anxiety-reducing and comforting effects. The concept of emotional eating originates from this theory and is developed further by Hilde Bruch (1961). In her article Bruch (1961) propounds that when individuals try to handle their negative emotional states through eating, they perform ‘emotional eating’.

According to various research, emotional eaters cannot distinguish hunger from the physiological effects of negative arousal states (Bruch, 1961; Konttinen et al., 2010; van Strien et al., 1986). Due to this confusion, which probably results from faulty early learning experiences, they initiate

emotional eating behavior (Bruch, 1961). In addition to the problematic learning relation, some individuals tend to utilize the eating activity as a distraction from their internal conflicts. Instead of struggling with the negative thoughts and emotions, they create an escape route by giving their attention to the currently available stimulus, which is the food itself (Ouwens et al., 2009). By performing this act, emotional eaters aim to generate a sense of temporary peace with oneself during the eating process.

As emotional eating started to attract attention from the researchers in the late 90s, new psychometric scales have emerged for its measurement. One of the well-known and valid scales for this purpose is The Dutch Eating Behavior Questionnaire (DEBQ) developed by van Strien et al. (1986). The authors attributed the performance of overeating into three different factors; psychosomatic eating (emotional eating), external eating, and restraint eating. Participants could indicate their answers using a five point Likert scale; (1) indicating never, (5) indicating very often and 13 out of 33 questions were specifically measuring the emotional eating factor. However, it appeared to us, that the themes of these 13 questions rather focused on the concept of coping with the negative emotions rooting from negative arousal states such as anxiety, distress or rejection. We found this scale to be incomprehensive regarding other possible and uninvestigated roots for awaking negative internal states such as complying to others in social situations.

Almost a decade later, Arnow, Kenardy, & Agras (1995) developed a new scale called Emotional Eating Scale (EES), which measures emotional eating directly and specifically. The five-point Likert-type scale had 25 items each indicating a negative emotion and measuring the relative eating

urge. While EES addresses a wide variety of emotions, we would like to criticize this scale for the wording of its items. The authors designed the items in a single-word format such as “Resentful”, “Discouraged”, and “Jealous”. Similarly, the test takers were also expected to score their eating urge regarding the emotion defined in a single word. We argue that addressing the test takers with this type of direct, context-free items would hinder them to relate with those emotions during the test. Therefore, we think that this scale placed little importance on the context of the emotions’ emergence, but focused on the emotion itself.

Jackson et al. (2003) interpreted the eating -and drinking- behavior and concluded that it can be initiated by four different motivations; coping, social, compliance and pleasure related motivations. Coping as a motivation comes into play as the individual faces with negative arousal states, as it is the case in the previously mentioned scale examples. Social motivations arise when the individual is presently in a social gathering and is expected to eat as an indication of enjoyment and celebration. Similarly, compliance-related eating is observed when the individual is expected and also pressured to eat in a social context. Lastly, pleasure-related motivations are observed in individuals who enjoy the act of eating or perform it as a reward. After deciding on these four factors, the authors developed the Motivations to Eat Scale, which consisted of 39 items with 17 of them on the coping, 7 of them on the social, 7 of them on the compliance and 8 of them on the pleasure subscale. When examining this scale, we noticed contextual similarities between social and compliance items and therefore concluded that they could have been gathered under a single factor.

Rationale

In the light of the information presented above, this study aims to develop an emotional eating scale comprising of two factors; coping and compliance. We desired to focus on the eating behavior mainly emerging from the negative arousal states and therefore, we eliminated any data collection about the positive states such as eating for a reward, pleasure, and happiness. Additional to these two emotion-related factors, we decided to collect data about another construct, the act of eating without being hungry, which theoretically stands for emotional eating itself. Thus, in this current study the subjects were given a Qualtrics survey, during which they were asked to score given situational items in either five-point Likert scales as (1) indicating never and (5) indicating very often/usually or two-point true (1)/ false (2) questions. The majority of the sample at hand consist of Koç University undergraduate students, but any healthy Turkish adult with the age ranging from 18 to 65 were eligible to take the survey.

Methods

Participants

The final sample consisted of 162 Turkish individuals aged between 18 and 65. All the participants choose to attend the survey freely and without seeking any profit for their attendance.

Materials

Our scale consists of 19 questions having two proposed factors of emotional eating: the effect of societal pressures and friends (i.e. compliance), using eating as a means to deal with psychological stressors (i.e. coping). It additionally has a third variable: the physical state of the individual (i.e. physiology). It is important to note that the third variable is technically not a factor of emotional eating, as physiological eating indicates non-emotional eating. Nonetheless,

it is included to see how the physical state of hunger influences emotional eating. Five items were generated for compliance (e.g., “I tend to want to eat if my friends are eating as well”), and seven items were generated for coping (e.g., “I feel the need to eat when I am lonely”) and physiological (e.g., “I stop eating when I feel full”) each, making 19 items in total.

Our group members translated the short form of Marlowe Crowne’s Social Desirability Scale (MCSDS) (Reynolds, 1982) to Turkish. MCSDS measures the test taker’s proneness to display social desirability and assesses if the respondents misrepresent themselves to manage their self-presentation. MCSDS demonstrates an acceptable level of reliability ($r = .76$). The scale has 13 items in true/false response format, encoded as 1 & 2 respectively. We predicted this scale to be related with compliance and hence used it as one of the validation scales.

The other validation scale was the Turkish adaptation of Perceived Stress Scale (PSS) (Eskin et al., 2013) with an internal consistency of $r = 0.84$ and a test-retest reliability coefficient of $r = 0.87$. The five-point Likert scale consisted of the options: never, almost never, sometimes, often, very often encoded as 1, 2, 3, 4, 5 respectively.

The final validation scale that we used was the Turkish adaptation of Mindful Eating Questionnaire (MEQ) (Köse et al., n.d.) with a Cronbach’s alpha of 0.73. The five-point Likert scale consisted of the options: never, rarely, sometimes, often, usually, encoded as 1, 2, 3, 4, 5 respectively.

Procedure

The Qualtrics link for the tests was distributed to the participants via social media platforms such as WhatsApp and Facebook. All the participants were presented the tests in the following order: BEEM scale, Marlowe Crowne’s Social Desirability Scale, Perceived

Stress Scale and Mindful Eating Questionnaire. For five consequent days, data were collected from willing participants. Convergent validity and internal reliability test analyses were conducted.

RESULTS

Factor Analysis

Factor analysis was conducted on the generated scale, based on Eigen value being greater than one yielded six factors. So, the analysis was repeated by fixing the number of factors to three. When investigating the item loading on these three factors it was observed that items 1 (P1), 4 (P2) and 10 (P4) were not loading on one factor alone and thus were eliminated. The remaining 16 items yielded three factors with the extraction sum of squared loadings of 51.505 percent, with a KMO score of .831, χ^2 (1, N = 120) = 905.085, $p < .001$.

Reliability Analysis

Reliability analysis for the 16 items yielded a Cronbach's alpha of .60. A split-half reliability analysis resulted in a Spearman-Brown coefficient of .86, meaning the test was reliable.

Correlation

To conduct correlation analysis, the mean value for the proposed factors of emotional eating (Coping & Compliance), the mean value for the variable (Physiology) and the mean value for the three previously generated scales (Marlowe Crowne's Social Desirability Scale, Perceived Stress Scale and Mindful Eating Questionnaire) were computed and a Pearson correlation analysis was conducted using these data. As a result, there was a correlation between coping and perceived stress, $r(162) = .314$, $p < .01$, between physiology and mindful eating, $r(162) = .607$, $p < .01$, between coping and mindful eating, $r(162) = .808$, $p < .01$ and between compliance and mindful eating, $r(162) = .331$, $p < .01$. However, there was no

correlation between compliance and social desirability, $r(162) = -.110$, $p > .05$. Additionally, a correlation was observed between all three generated variables, coping and compliance, $r(162) = .339$, $p < .01$, coping and physiology, $r(162) = .492$, $p < .01$ and compliance and physiology, $r(162) = .375$, $p < .01$.

Discussion

In this study we aimed to develop a brief and convenient scale to measure Turkish people's emotional eating habits. The Cronbach's alpha of our 16-item scale is found to be .60, which is moderately high considering our low item count. The participants were not instructed to fill the survey at any particular time of the day; thus, the hunger level of the participants might be different. This variance might be a reason lowering the reliability of our scale. To increase reliability, further studies should be done where the scale is administered to subjects with similar hunger levels, perhaps by allowing the test to be completed only after a certain time has passed after a meal.

When generating our scale, we have separated the construct "emotional eating" into two motivational factors as coping and compliance. Additional to these two, we have conceived of a third factor, which was used to discover the physiological eating habits of our test takers. As mentioned above, the third factor is not a motivational factor, and hence, it (is theoretically not a factor of emotional eating as the other two) does not take a theoretical part in our emotional eating construct. However, during factor analysis, we fixed the number of factors to three since these three variables differ from each other on a theoretical basis. Even though the third factor concerning physiology is not a factor of emotional eating, we included it in the factor

analysis to see its possible effects on the other two factors.

Our analysis showed significant correlation between all three generated aspects of emotional eating. This might indicate that our items did not exclusively load on one factor only. The items could have been chosen more carefully to only correspond to one factor alone. The three factors explained 51.5% of the total variance, which again might show that our factor loadings were not very precise.

Our analyses showed significant convergent correlation between the proposed factor of coping and the Perceived Stress Scale (PSS) and between physiology and the Mindful Eating Questionnaire. (MEQ) However no correlation was found between compliance and Marlowe-Crowne Social Desirability Scale (MCSDS) as we had hypothesized. Perhaps its reliability was affected because the scale was a translation and not an adaptation. These results show that PSS was an adequate analog scale to evaluate coping and MEQ was adequate to evaluate physiology. However, a different scale from the literature should have been used to test compliance as no correlation was found. As stated an adaption of the MCSDS could have been used instead.

Our study is not free of limitations: firstly, our sample consists of only non-clinical participants. The usage of clinical participants, meaning clinically over- or underweight individuals, is necessary to understand the applicability of this scale to those groups. Secondly, eating is a culture-bound phenomenon, and therefore our scale might be less applicable to minority groups or people from other cultures. For a better understanding of the scientific and practical value of the scale and for obtaining culturally fair results, it is essential to translate and adapt the scale to other languages.

Among the previously established scales, the Mindful Eating Questionnaire adapted to Turkish by Köse et al (n.d.) was a study conducted at a university and was not published in an academic journal, although our items showed convergent validity, the validity of the scale itself may be questionable.

Since being overweight is a common problem in the developed world, understanding what drives one to overeating is a key area of research. Improvements of this scale and its usage in future studies might significantly contribute to the literature on emotional eating by augmenting the number and theoretical scope of the factors. We hypothesized that “pleasure” would also be a fit for a motivational factor of the emotional eating construct. Since it is a positive internal state, we did not elaborately examine its effect on appetite, but it seems to be a promising research topic.

In conclusion, the BEEM emotional eating scale is a valid and reliable scale that can be utilized to discover the nature of one’s eating patterns, to see if one eats out of hunger, or if other psychological factors affect the eating behavior. The factors that can be outlined by this scale are eating for coping with personal distress and eating for complying to others.

References

- Arnou, B., Kenardy, J., & Agras, W. S. (1995). The Emotional Eating Scale: the development of a measure to assess coping with negative affect by eating. *The International Journal of Eating Disorders, 18*(1), 79–90.
- Bruch, H. (1961). Transformation of oral impulses in eating disorders: a conceptual approach. *The Psychiatric Quarterly, 35*, 458–481.

- Heatherton, T. F., Peter Herman, C., & Polivy, J. (1991). Effects of physical threat and ego threat on eating behavior. *Journal of Personality and Social Psychology*, *60*(1), 138–143.
- Eskin, M., Harlak, H., Demirkıran, F., & Dereboy, Ç. (2013). Algılanan Stres Ölçeğinin Türkçeye Uyarlanması: Güvenirlik ve Geçerlik Analizi. *The Adaptation of the Perceived Stress Scale Into Turkish: A Reliability and Validity Analysis.*, *51*(3), 132–140.
- Jackson, B., Lynne Cooper, M., Mintz, L., & Albino, A. (2003). Motivations to eat: Scale development and validation. *Journal of Research in Personality*, *37*(4), 297–318.
- Kaplan, H. I., & Kaplan, H. S. (1957). The psychosomatic concept of obesity. *The Journal of Nervous and Mental Disease*, *125*(2), 181–201.
- Konttinen, H., Männistö, S., Sarlio-Lähteenkorva, S., Silventoinen, K., & Haukkala, A. (2010). Emotional eating, depressive symptoms and self-reported food consumption. A population-based study. *Appetite*, *54*(3), 473–479.
- Köse, G., Tayfur, M., Birinciöglu, İ., & Dönmez, A. (n.d.). Yeme Farkındalığı Ölçeği'ni Türkçeye Uyarlama Çalışması. Retrieved from <http://www.ejmanager.com/mnstemps/77/77-1480327096.pdf>
- Ouwens, M. A., van Strien, T., & van Leeuwe, J. F. J. (2009). Possible pathways between depression, emotional and external eating. A structural equation model. *Appetite*, *53*(2), 245–248.
- Reynolds, W. M. (1982). Development of reliable and valid short forms of the marlowe-crowne social desirability scale. *Journal of Clinical Psychology*, *38*(1), 119–125. [https://doi.org/10.1002/1097-4679\(198201\)38:1<119::AID-JCLP2270380118>3.0.CO;2-I](https://doi.org/10.1002/1097-4679(198201)38:1<119::AID-JCLP2270380118>3.0.CO;2-I)
- van Strien, T., Frijters, J. E. R., Bergers, G. P. A., & Defares, P. B. (1986). The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *The International Journal of Eating Disorders*, *5*(2), 295–315.

Development of Premarital Sex Attitude Scale

Ekin Çakır, Ezgi Yıldız, Dilara Alkan, Meryem Giden, Bengi Türer

Middle East Technical University

In this study, the validity and reliability of the recently developed Premarital Sex Attitude Scale (PSAS) were assessed. The study is conducted in Middle East Technical University with 40 participants. Results indicated that the scale is a valid one and has a .95 inter-item reliability. Moreover, participants who had higher scores in PSAS were found to have positive attitudes towards having premarital sex. Their religiosity score was low; this score was used to assess convergent validity. To assess divergent validity, Openness to Experience Scale was used and as expected, the correlation with the PSAS was not significant. No gender difference was assessed.

Keywords: premarital sex, religiosity, validity, Premarital Sex Attitude Scale, Openness to Experience Scale

In this study, the concept of premarital sex is defined as involvement in any act of sexual behavior and physical closeness including touching, kissing, sexual intimacy without intercourse and sexual intercourse. People's attitudes toward premarital sex can be affected by their religious beliefs, economic development of the country they live in (Jung, 2016) and culture (Sprecher, & Hatfield, 1996). For example, Asian cultures are perceived to be more collectivistic. Thus, it is likely that individuals' attitudes and behaviors toward premarital sex will be affected by familial and societal values. On that note, it is also reasonable to expect that their sexual attitudes will be interiorized within the frame of marriage (Jaafar, Wibowo, & Afiatin, 2006). The current study aimed to develop a psychological test that assesses attitudes of

Turkish people towards premarital sex. The study presents its factorial properties by comparing its validity with other psychological constructs, namely religiosity and openness to experience.

The concept of premarital sex was tested in many studies, but a clear-cut definition of it seems to be lacking. Some premarital sexual studies focused on societal and religious effects on attitudes toward premarital sex in different cultures (Hojat, Shariat, & Ansar, 2015; Jung, 2016). On the other hand, some studies concentrated on the attitude changes toward premarital sex over time. One reason of the changes in attitudes toward premarital sex is the changes resulting from cohort effects. For example, recently, people have reduced attendance at religious activities and had more access to education, causing change in their

attitudes. According to another study, other reasons of those changes except for cohort effect are general trends, period and age effects. People have been more permissive toward premarital sex since early 1960s, which was the general trend within that time (Harding, & Jencks, 2003). In addition, examples around the world (e.g. Thailand) suggest that the attitudes toward premarital sex can also change with the influence of Western culture (Rasamimari, Dancy, Talashek, & Park, 2007). This concept was also investigated in terms of parental attitudes and it was found that individuals, who have high permissive parents toward premarital sex, have more permissive attitudes as well (Gravel, et al., 2016).

Based on the literature review, one of the oldest and most widely used scales to measure the premarital sex attitude was developed by Reiss in 1964 by adapting Guttman's scale. This scale measures to what extent people are open to engage in sexual behavior and how this attitude differs for women and men. The measure consists of two subscales each with 12 items separately for men and women. It was constructed based on the American culture in order to assess Americans attitudes. The items question participants' attitudes about premarital sex by combining the four levels of affection, which are love, engagement, strong affection and no affection in the relationship. These levels of affection were interrelated to three physical behaviors: kissing, touching and sexual intercourse. The adaptability coefficient for that scale was found to be over .90 for each subscale when compared based on Guttman's scale which has a validity of .90 (Hampe, Howard, & Ruppel, 1974). However, Reiss' scale has some limitations. Firstly, it is a bit outdated to measure attitudes

of new generations and it is only limited to the American culture. Secondly, Reiss' scale remains inadequate to address more varying degrees of sexual behavior; rather it only contains three types of behavior (Clayton, 1979). Finally, it is better to regard the concept of relationship as evolving in time and having certain stages ranging from first dating to engagement, rather than specifying only four levels of affection, which pretty much overlap with each other (Sprecher, McKinney, Walsh, & Anderson, 1988). Considering these limitations, the test needs to be updated and adapted according to cultures' sexual norms. The types of behavior can be enlarged to include other behaviors such as oral sex and petting.

The other scale is Attitudes toward Sexuality Scale (ATSS) that was developed by Fisher (1988). Items are remodeled from Calderwood's Checklist of Attitudes toward Human Sexuality (1971). ATSS measures attitudes toward sexuality in three different stages of adolescents (early, middle and late) and compare their results with their parents. ATSS consists of 13 items and participants are provided with a 5-point Likert scale in order to indicate their answers. Two different concepts are measured in ATSS, which are sexually liberal and sexually conservative orientations. The Cronbach's alpha value of that scale for the entire population is .80 (Fisher, & Hall, 1988). However, ATSS has some shortcomings as well. First of all, the language of the scale is more applicable for middle class participants than lower class making it hard for individuals from other economical classes to understand it (Fisher, & Hall, 1988). Secondly, that scale may not be useful for measuring general sexual attitudes because it does not contain some

important concepts related to sexual attitudes such as masturbation and oral sex.

As can be seen in previous literature examples, premarital sex was investigated many times before. While investigating premarital sex, religiosity has been a much-studied concept in terms of its relationship with sexual attitudes and behaviors (Peterson, & Donnenwerth, 1997). It has been seen by social scientists that religious factors are highly correlated with individual's specific moral orientations towards an array of social, political, and cultural issues that involve relationships and family (as cited in Ogland, & Hinojosa, 2012). On that note, recent evidence suggests that religion, in fact, has a powerful influence on premarital sex attitude: People who have a strong belief in religion and join religious rituals have been found to be more likely to disagree more with permissive premarital sex attitude (Jung, 2016). Moreover, according to the Social Learning Theory, due to religions' prohibitory stance over premarital sex, individuals who are exposed to their religion's norms adopt a negative attitude toward premarital sex (Akers, 1998). In the light of these information, religiosity was used as convergent validity in the current study.

Openness to experience is one of the personality traits connected with emotional, cognitive and motivational fields, and mostly related to creativity (McCrae, 1996). People with high openness feel the need to search for new experiences (Costa, & McCrae, 1992). To our knowledge, there is no study in the literature investigating the relationship between openness to experience and premarital sexual attitude. However, it seems reasonable to expect from people who are open to new experiences to have positive attitudes towards

premarital sex as well. Hence, to ensure that our test did not measure people's openness to new experiences and, as intended, measured people's attitudes towards premarital sex, Openness to Experience Scale was used to insure the divergent validity of our test.

As stated earlier, the main purpose of this study is to develop a scale, which can measure people's attitudes towards premarital sex as a general concept. Since this concept is one that shows great cultural variance, we think that PSAS being Turkish alone and targeting individuals in Turkey will be an important contribution to the literature. Furthermore, although some items like "I think TV shows that normalize premarital sex do not harm society's values." or "I believe, in schools there should be educational programs in which students are taught that premarital sex is something to be avoided." in the scale might seem irrelevant to some cultures, given the conservative nature of Turkey these items point to important issues. Carkoğlu and Kalaycıoğlu showed the increase in Turkey's conservatism as a result of religion by drawing attention to the increase in the number of Quran courses and the students joining to these courses (as cited by Göksel, 2013). Secondly, contrary to many other scales, which intend to measure the same concept, it was not designed to assess gender differences in attitudes towards premarital sex. It was rather developed considering the attitudes in a holistic manner. Other scales in the literature are mostly interested in attitudes about individuals' own sexuality. However, we thought that one's own sexuality is only one aspect of the general attitude. There are some other aspects and concepts, which form an overall attitude. Thus, we determined four different aspects as four

subscales. If Turkey is considered as a collectivistic culture (Hofstede Center, 2014), the attitudes of people in Turkish culture might be affected by those collectivistic value systems. Therefore, we took in consideration individuals' attitudes toward other people around them. By doing that, we investigated individuals' attitudes in a broader perspective than other scales. As a first subscale, we measured the attitudes of individuals toward sexuality of close ones, since a person might have a different perception of his/her close ones' sexual attitudes than his/her own sexual attitudes. The other aspect what we measured was perception of sexuality in the community dimension. Society's effect on attitudes toward premarital sex was regarded in that subscale. Additionally, we measured whether individuals consider premarital sexuality when they choose their partners. The last aspect is one's own sexuality. As a result, PSAS provided a multidimensional and overall attitude rather than a narrowed one. The third contribution of our scale is that contrary to other scales PSAS does not use intimacy as a sub factor (Reiss, 1964). PSAS includes several items with different levels of intimacy. For instance, "It is easy for me to have sexual intimacy with my romantic partner", "I am comfortable to have intercourse with my fiancée" and "I have sexual partner/s that I regularly meet without getting attached emotionally". However, we think that premarital sex attitude should not be explained differently based on presence or absence of intimacy. Instead, intimacy should be considered as a factor, which determines the degree of positiveness of attitudes. For instance, one might have a positive attitude about having a sexual relationship with his/her fiancé, but might be against having a sexual relationship

without an intimacy. Such a person can still be considered to have a positive attitude towards premarital sex. However, another person who thinks having a sexual relationship with both a fiancé and with someone who the person does not have any romantic intimacy has definitely a more positive attitude than the aforementioned person. Thus, intimacy's role is restricted to determining the level of positiveness.

In accordance with the literature, we determined four subscales, which can measure premarital sex attitudes in a holistic way and from a broader perspective. In the process of determining subscales, we observed that in Turkish culture there are some different aspects, which form a general attitude towards premarital sex. These aspects are a) one's own sexuality, b) sexuality of close ones, c) sexual attitudes in the community dimension, and d) sexuality in partner selection. After determining the subscales, based on our past experiences and observations, we wrote items that can point out important issues. While developing the items, we observed that our items were shaped by the Socialization Theory. According to this theory, people's learning processes are in tune with their culture and they live in compliance with it (Macionis, Gerber, John, & Linda, 2010). After creating an item pool, we rated our subscales to determine how many items each subscale would include. Based on our observations, we decided how many items would go under each subscale considering their degree of decisiveness about the general attitude towards premarital sex. In the process of selecting items from the item pool, we tried to eliminate the items, which measure similar concepts. We especially chose items, which represent important aspects of Turkish culture.

We hypothesized that: a) PSAS will have high negative correlation with Religiosity Scale in terms of convergent validity, b) PSAS will not be relevant to Openness to Experience Scale in terms of divergent validity, c) Individuals whose scores are high in PSAS will score higher in the criterion related validity scale.

Method

Participants

40 students (20 women, 20 men) from different departments of the Middle East Technical University participated in the study, by using convenience sampling. Participants were between the ages of 19 and 25 ($M = 21.45$, $SD = 1.47$). Participants' gender, sexual orientation or their relationship statuses were not considered.

Instruments

Demographic Questions. There were 9 demographic questions including gender, age, the city they grew up, which department they are in, where they are living right now (dormitory, in separate house or with family) etc. The question about gender was not interested in individuals' biological sexes. It was an open question for participants to describe themselves in the way they feel.

Premarital Sex Attitude Scale (PSAS). This scale is developed by the authors to measure the premarital sex attitude of Turkish people. It is a 5-point Likert type scale from 1- strongly disagree to 5- strongly agree. It has 24 items, four reversed, and can be divided into four subscales, namely a) attitudes towards one's own sexuality (It is okay for me to have physical intimacy with my partner without having intercourse.), b) attitudes towards sexuality of close ones (It does not matter to me

if a potential friend had premarital sex or not.), c) attitudes towards sexual attitudes in the community dimension (I think people need to have premarital sex to see if there is sexual harmony.), and d) importance of sexuality in partner selection (Finding out that my partner, to whom I am seriously considering to marry, had sex with another person before would not affect my decision to get married with him/her.). Also, to set an example of a reverse item: "I am against premarital sex because of my respect for my potential partner." The maximum obtainable score from that scale is 120 and the minimum is 24. Participants who got scores above the mean plus standard deviation were considered to have positive attitudes towards premarital sex. Also, to ensure that there was not any random respondent, we included filler item (If you read this question please mark 3). This item was not considered in statistical analysis.

Religiosity Scale. To measure convergent validity, the Turkish adaptation of Peterson and Seligman's (2004) Religiosity Scale, was used (Demir, & Kumkale, 2013). It is a 7-point Likert type scale from 1-strongly disagree to 7-strongly agree, and has 10 items, one reversed item. It has items such as "I think my beliefs make my life important." to measure participants' religious tendencies. It is a valid and reliable scale with a Cronbach's alpha of .90.

Openness to Experience Scale. To measure the divergent validity, Openness to Experience Scale was used from the Big Five Inventory. It is a 5-point Likert type scale from 1-strongly disagree to 5-strongly agree. It has 10 items among which two are reversed, which can be exemplified as "I see myself as someone who is original, comes up with new ideas." It is also a

valid and reliable test, which was adapted to Turkish by Karaman, Doğan, and Çoban (2010). The test was originally developed by John, Donahue, and Kentle (1991) and has a Cronbach's alpha of .86.

Criterion Validity Questions. To measure criterion validity, a 5-question scale was developed by the authors. The questions were related to participants' past relationships, past sexual activities and their opinion about living with a romantic partner. The scale does not have reversed items and all questions were either open ended or yes-no questions. For all answers, a point was assigned in order to calculate the scores. For example, for the question "How many girlfriends/boyfriends did you have?" if the answer was 0, the score for that item would be 0; if the answer was 1 or two 2, the score would be 1; and if the answer was 3 or more, the score would be 2.

Procedure

The ethical procedure was followed. The informed consents and the scales were collected separately from participants and these were put into different envelopes for the sake of anonymity. A day after participants received the scale package consisting of demographic questions, PSAS, Religiosity Scale (Demir, & Kumkale, 2013), Openness to experience items in Big Five Inventory (Karaman, Doğan, & Çoban, 2010) and the criterion questions. In order to make the participants feel comfortable, they were asked not to sign the informed consent but write only their initials since the questions were about their private lives. After the collection of scales, participants were given a debriefing form about the study's details and its purpose.

Results

Descriptive statistics of the premarital sex attitude scale items were calculated. Results of the PSAS were $M = 3,67$ and the $SD = .83$, with a minimum score of 1.42 and a maximum score of 4.83. In addition, Religiosity scale items were averaged, and the result was found as $M = 3.78$ and $SD = 1.71$. For Openness to Experience dimension in Big Five Inventory, it was found as $M = 4.00$ and the $SD = .36$. Mean of the criterion-related questions' was .81 and its SD was .44.

Reliability of the premarital sex attitude scale was measured by using Cronbach's alpha. Internal consistency reliability of the premarital sex attitude scale was .95 with an alpha level of .01. According to item-total statistics, there would be no significant change in the Cronbach's alpha in the case of deleting any item. Therefore, no item was deleted (see Table 3). In addition, as can be seen from Table 2, reliability values of the importance of sexuality in partner selection, attitudes toward one's own sexuality, attitudes toward sexuality of close ones, attitudes toward sexual attitudes in the community dimension sub factors was measured (respectively, .89; .89; .78; .85).

The correlation between religiosity and the Premarital Sex Attitude Scale was measured for the analysis of the convergent validity. Results showed that there is a significant high negative correlation between religiosity and PSAS ($r = -.71, p < .01$). In addition, sub factors of the Premarital Sex Attitude Scale, which are attitudes toward one's own sexuality, attitudes toward sexuality of close ones, attitudes toward sexual attitudes in the community dimension, and importance of sexuality in partner selection, showed

significant negative correlation with religiosity (respectively, $r = -.59$; $-.69$; $-.71$; $-.66$).

Openness to Experience in the Big Five Inventory was used to measure divergent validity, which displayed no significant correlation with the Premarital Sex Attitude Scale. ($r = .26$, $p > .05$). Only the attitudes toward one's own sexuality sub factor showed significant correlation with openness to experience ($r = .34$, $p < .05$) (see Table 4).

Five questions were developed to assess the criterion-related validity and displayed a moderate correlation with the Premarital Sex Attitude Scale ($r = .45$, $p < .01$). However, for the sub factor "attitudes toward sexuality of close ones", no significant correlation was found with criterion-related questions ($r = .28$, $p > .05$).

Discussion

The aim of the present study was to develop a new scale with four subscales that can measure the premarital sex attitudes in a holistic manner. The scale consists of 24 items and has high internal consistency reliability. Since no item would make a significant difference in reliability if deleted, all items were kept in the scale.

In order to validate our results, we investigated correlations. For convergent validity, as we had expected, a strong negative correlation was found between Religiosity Scale and PSAS. This finding indicates that participants who scored higher in religiosity had more negative attitudes towards premarital sex. For divergent validity, we had expected to find no significant correlation between Openness to Experience and PSAS. Results supported our hypothesis. A moderate or high

correlation had been expected for criterion validity and a moderate correlation was found between 5 criterion questions and PSAS.

The PSAS was developed considering the elements of the Turkish culture. Thus, it can be applied to people who live in Turkey. Even though Turkey has been being considered a European country, the culture and the attitudes towards certain issues have not changed much. For instance, while premarital sex is not a serious problem in Western cultures, people still could face serious problems due to premarital sex in Turkey. Between the years of 2010 and 2016, 90 women fell victim to honor crimes and murdered by their husbands, family members, relatives and other acquaintances (Kadın Cinayetlerini Durduracağız Platformu, 2017). Therefore, premarital sex attitudes are a critical issue for social psychologists and sociologists to investigate. PSAS can be used in these investigations in order to detect the factors that cause and affect such positive or negative attitudes, and what these factors depend on. Since PSAS has items related to sexuality in the community, it would be useful for social and cultural studies as well.

Although hypotheses about PSAS were supported by statistical analyses, there are some remarkable points. From the four sub factors, it was found that the sub factor of sexual attitudes in the community dimension is the most correlated factor to religiosity. Even though there may be certain mediators, it can be argued that religiosity has a strong effect on positiveness of the attitudes towards premarital sex, especially in terms of sexual attitudes in the community dimension. The second important point is that PSAS predicted the behaviors, which were mentioned in criterion questions at a moderate level. It can be inferred

that at a moderate level, people who have more positive attitudes towards premarital sex are more likely to cohabit and have sexual relationship before marriage. The frequencies of our data showed that the range was wide both in PSAS and Religiosity Scales. Thus, even though our sample size was small almost all participants had different attitudes and different levels of religiosity.

One limitation of the present study is that there were only 40 participants. This small sample size could create a problem in terms of the representative qualities of our study. Also, since all participants were METU students, who were a young population and were relatively higher in education, the population under investigation represents only a small proportion of Turkey. A problem related to this is about the Openness to Experience scale. Although there is a small correlation, Openness to Experience scores of almost all participants were really high with a mean of 4.00 in a 5-point Likert type scale. This may have two reasons. The first one is that all participants were university students. Individuals may tend to be open to new experiences in these young ages. Also, being in a university may make them more open to new experiences. The second possible reason is social desirability. Since the questions of the Openness to Experience scale can be considered to define an ideal social university student or a popular person in a group of friends, participants might have tended to answer them positively. Another limitation is about the relationship between religiosity and premarital sex attitudes. Even though participants had different levels of religiosity, there are some important points to consider. First of all, there may be religion related and denominational differences in terms

of religious rules and flexibilities. Since the Religiosity Scale is not interested in religious practices that religion requires, individuals with less restricted religious rules about premarital sex might not be differentiated. The second one is that we observed that in the “I try to perform my religion’s requirements” item, some religious people chose “I strongly disagree”. Not only whether people are religious or not, but also whether people think that they should perform the religion’s requirements is very important. Even if a person is religious, if he/she does not think that performing the requirements of the religion is necessary or important, then the effect of religion on attitudes towards premarital sex weakens. Another limitation is that the age range of the participants might have affected the way they answered the criterion questions. Someone who has positive attitudes towards premarital sex may not have had a sexual relationship due to his/her young age. Likewise, young people may not want to cohabit with their partners in such a young age. They may prefer to live with their friends or they may simply not want to develop a serious and strong bond with their partners.

Considering the limitations, future research should conduct a study including individuals from different age levels as well as different educational backgrounds. Using a large sample, the relationship between premarital sex attitudes and some demographic properties could be investigated. The items of Openness to Experience scale are measuring openness to experience in an overt manner, and this may cause social desirability and lead to false answers given by individuals. Moreover, individuals may have different standards in terms of rating themselves due to their self-perceptions. Thus, in order to measure it more

realistically, a covert scale may be used for future studies. For instance, “How many art galleries have you visited in the last 6 months?” or “How many art magazines are you following?” can be better indicators in this case than the item “I have little interest to art”. By using a covert scale, the possible social desirability bias can be decreased. Furthermore, since the Religiosity Scale does not differentiate people who follow the rules of religion with the ones who do not; another scale that measures this difference as well can be used. Moreover, in order to identify the reason of strong effect of religiosity on premarital sex attitudes, quasi experiments can be conducted.

The current study reveals that individuals with high religiosity have negative attitudes towards premarital sex; openness to experience level is not relevant to attitudes towards premarital sex; individuals who have premarital sex have positive attitudes towards premarital sex; and PSAS can measure what it aims to measure. PSAS is expected to be used in cultural and social studies since it investigates how different aspects shape individuals’ general attitudes.

References

- Akers, R. L. 1998. *Social Learning and Social Structure: A General Theory of Crime and Deviance*. Boston, MA: Northeastern University Press.
- Clayton, R., R., Bokemeir, J. L. (1980). Premarital sex in seventies. *Journal of Marriage and the Family*, 42(4), 759-775. doi: 10.2307/351823.
- Costa, P. T., & McCrae, R. R. (1992). Revised NEO personality inventory and NEO five-factor inventory: Professional manual. Odessa, FL: Psychological Assessment Resources.
- Demir, B., & Kumkale, G. T. (2013). Individual differences in willingness to become an organ donor: A decision tree approach to reasoned action. *Personality and Individual Differences*, 55(1), 63-69. doi: 10.1016/j.paid.2013.02.002.
- Ellias, V. L., Fullerton, A. S., & Simpson, J. M. (2015). Long-Term Changes in Attitudes Toward Premarital Sex in the United States: Reexamining the Role of Cohort Replacement. *The Journal of Sex Research*, 52(2), 129-139. doi: 10.1080/00224499.2013.798610.
- Fisher, T. D., & Hall, R. G. (1998). A scale for the comparison of the sexual attitudes of adolescents and their parents. *Journal of Sex Research*, 24(1), 90-100. doi: 10.1080/00224498809551400.
- Gravel, E. E., Young, M. Y., Darzi, C. M., Olavarria-Turner, M., & Ming-See, L. (2016). Premarital sexual debut in emerging adults of South Asian descent: The role of parental sexual socialization and sexual attitudes. *Sexual & Culture*, 20(4), 862-878. doi: 10.1007/s12119-016-9362-1.
- Göksel, İ. (2013). Female labor force participation in Turkey: The role of conservatism. *Women’s Studies International Studies*, 41 (1), 45-54. Doi: 10.1016/j.wsif.2013.04.006
- Hampe, G. D., Ruppel, J. R., & Howard, J. (1974). The measurement of premarital sexual permissiveness: A comparison of two Guttman scales. *Journal of*

- Marriage and the Family*, 36(3), 451-463. doi: 10.2307/350716.
- Harding, D. J., & Jencks, C. (2003). Changing Attitudes toward Premarital Sex. *Public Opinion Quarterly*, 67(2), 211-226. doi:10.1086/374399.
- Hofstede Center (2014). Scores for individualism for Turkey. Retrieved from <http://geert-hofstede.com/turkey.html>
- Hojat, M., Shariat, S. V., & Ansari, S. (2015). Iranian college students' attitudes toward premarital sex, marriage, and family before and after the Islamic Revolution. *International Journal of Sexual Health*, 27(3), 224-234. doi: 10.1080/19317611.2014.957794.
- Jaafar, J., Wibowo, I., & Afiatin, T. (2006). The relationship between religiosity, youth culture, and premarital sex among Malaysian and Indonesian adolescents. *Asia Pacific Journal of Social Work and Development*, 16(2), 5-18. doi:10.1080/21650993.2006.9755999.
- Jung, J. H. (2016). A cross national analysis of religion and attitudes toward premarital sex: Do economic context matter? *Sociological Perspectives*, 59(4), 798-817. doi: 10.1177/0731121415595428.
- Kadın Cinayetlerini Durduracağız Platformu (2017). Retrieved from <http://kadincinayetleri.org/>
- Karaman, N. G., Doğan, T., & Çoban, A. E. (2010). A study to adapt the Big Five Inventory to Turkish. *Social and Behavioral Sciences*, 2(2010), 2357-2359. doi: 10.1016/j.sbspro.2010.03.336.
- Macionis, Gerber, John, Linda (2010). *Sociology 7th Canadian Ed.* Toronto, Ontario: Pearson Canada Inc. p. 109.
- Martel, L. D., Hawk, S., & Hatfield, E. (2004). Sexual behavior and culture. *Encyclopedia of Applied Psychology*, 385-391. doi:10.1016/b0-12-657410-3/00210-5.
- McCrae, R. R. (1996). Social consequences of experiential openness. *Psychological Bulletin*, 120, 323-337. doi:10.1037/0033-2909.120.3.323.
- Ogland, C. P., & Hinojosa, P. (2012). Religion and social attitudes: Examining the contour of religion in moral judgments toward premarital sex and cohabitation in contemporary Brazil. *Sociology of Religion*, 73(4), 411-428.
- Petersen, L. R., & Donnenwerth, G. V. (1997). Secularization and the influence of religion on beliefs about premarital sex. *Social Forces*, 75(3), 1071-1088.
- Rasamimari, A., Dancy, B., Talashek, M., & Park, C. G. (n.d.). Predictors of sexual behaviors among Thai young adults. *Journal of the Association of Nurses in AIDS Care*, 18(6), 13-21.
- Reiss, I. L. (1964). The scaling of premarital sexual permissiveness. *Journal of Marriage and the Family*, 26(2), 188-198. doi: 10.2307/349726.
- Sprecher, S., & Hatfield, E. (1996). Premarital sexual standards among U.S. college

students: Comparison with Russian and Japanese students. *Archieves of Sexual behavior*, 25 (3), 261-288. Doi: 10.1007/BF02438165

Reiss' Premarital Sexual Permissiveness Scale. *Journal of Marriage and the Family*, 50(3), 821-828. doi: 10.2307/352650.

Sprecher, S., McKinney, K., Walsh, R., & Anderson, C. (1988). A revision of the

An Insidious Stressor: The Concept of Internalized Homonegativity and its Relation to Ethnicity, Religious Settings, Sexual Orientation Disclosure and HIV Risk Behaviors

Mert Nedim Mercan

Yeditepe University

This review paper examines the effects of internalized homonegativity on sexual orientation disclosure and risky sexual behavior. After defining the concept of internalized homonegativity, some of the triggers of this stressor such as religion and ethnicity are examined with respect to the research that has been conducted. Moreover, how religion and ethnicity affect the level and the process of developing internalized homonegativity, sexual orientation disclosure and risky sexual behaviors are inspected. According to previous studies, when the level of internalized homonegativity is higher, that is, when a person is exposed to relatively more stigmatization and experiences negative feelings more intensively, they engage in riskier sexual behavior. This paper takes a Western perspective towards the relation between aforementioned concepts and behaviors, and emphasizes that more cross-cultural research is needed to fully understand the impact of internalized homonegativity.

Keywords: Internalized Homonegativity, Sexual Orientation Disclosure, HIV Risk Behaviors, Religious Settings, Ethnicity

Internalized homonegativity (IH) is a critical and influential factor in psychological and social well being of lesbians, gays, and bisexuals (LGB). Even though there are a couple of different definitions in literature, internalized homonegativity is generally defined as a process of internalization of the negative attitudes towards homosexuality (Barnes & Meyer, 2012). Frost and Meyer (2009) state that many identity development theories support the idea of internalized homonegativity (IH). The phenomenon of internalized homonegativity can develop in as early as the pre-genital stages of psychosexual development. According to Freud's

psychosexual development theory, in this stage children start to discover their sexual identity. Later in the end of latency and adolescence, they may experience homosexual desires, while some may engage in homosexual activities, others may suppress their desires due to external biases. The suppression affects their perception of homosexuality and has the potential to turn into a homophobic attitude in adulthood (Malyon, 1982). The internalization process is categorized in various ways such as identity, cognitive, emotional, and behavioral (Malyon, 1982). Later, Meyer (1995), one of the pioneers of the concept of IH, defined the

concept as the conscious or unconscious internalization process performed by LGB persons and of the negative, hostile, and discriminatory messages perceived by the society. The definition of internalized homonegativity may resemble homophobia; however, there is a difference between these two concepts on an emotional setting.

Homophobia is a type of negative attitude toward homosexuality associated with the core emotions of fear and anxiety. It is argued that internalized homonegativity is broader than homophobia because it centers on disgust and anger alongside with fear and anxiety (Herek, 2004). Negative feelings and attitudes like this sort are generally learned via normative socialization. Once they are internalized, homonegativity may appear as an active stressor any time in one's life (Barnes & Meyer, 2012; Meyer & Dean, 1998). It would be appropriate to separate developing internalized homonegativity from the experience of homonegativity. Development, as stated before, refers to the process of internalization and it can be higher or lower according to the external stigmatization that people face while experiencing it. In another words, the way that people deal with the external stigmatization is usually related to the self-stigmatization to their psychological state in daily life (Kelley & Robertson, 2008).

Previous studies claim that IH has various influencers and triggers, e.g. ethnicity and religion (Amola & Grimmet, 2015; Balsam, Molina, Beadnell, Simoni & Walters, 2011; Barnes & Meyer, 2012; Domanico & Crawford, 2000; Harris, Cook & Kashubeck-West, 2008; Moradi et al., 2010; Schuck & Liddle, 2001). Amola and Grimmet's study regarding ethnicity showed that internalized homophobia was higher among ethnic minorities (2015). Three main theories; Minority Stress Theory, Distal Minority

Stress, and The Greater Risk Perspective Theory are important to conceptualize the effect of being an ethnic minority on internalized homonegativity.

Minority Stress Theory conceptualized by Meyer (1995) suggests IH is a type of stress, which results from the juxtaposition of minority and majority values and a confrontation with social surrounding. It could be argued that LGB people in ethnic minorities face greater deal of conflict in their daily life and are under higher risk of experiencing internalized homonegativity. Heterosexism, which is a form of homonegativity, is significantly higher in racial/ethnic minorities (Buston & Hart, 2001; Moradi et al., 2010). LGB people of minorities face micro aggression, especially in their communities, work places and even in some LGB communities, which can be discriminative toward people of color.

Distal Minority Stress Theory can be defined as disproportion in discrimination toward bisexual and homosexual individuals (Pachankis et al., 2015), and it partially defines a different type of stress for ethnical and sexual minorities. Moradi (2010) has also noted that LGB people of color and Latino LGB people are treated as outsiders in both LGB communities and in their own ethnic communities. They were excluded from more than one part of the society throughout the United States history. This type is not dependable for one's own perceptions and feelings, antithetically; it is more about external stigmatization (Operario & Fiske, 2001). For instance, if a man has a romantic or/and sexual relationship with another man and identifies himself as bisexual, he may perceive external stigma as homosexual and he will most likely become vulnerable for stress, anxiety and experience internalized homonegativity. In other words, bisexual men may be seen as homosexual although they

also engage in sexual activities with women (Pachankis et al., 2015; Lick, Durso & Johnson, 2013)

Another link between being an ethnic minority and experiencing internalized homonegativity is established in the Greater Risk Perspective theory. This theory argues that heterosexist discourse is wider in minority communities since LGB identity is seen as a White set-up (Moradi et al., 2010). Scientists who work on LGB identity theories assume that defining one's identity by sexual orientation is a perspective which is brought by White Europeans to America and processed to the native culture by acculturation, which later turned into assimilation (Bridges, Selvidge, & Matthews, 2003; Malebranche, Fields, Bryant, & Harper, 2009). Various studies about minority effect on internalized homonegativity suggest that the customs and gender roles of the minority culture (e.g. Latino and Black) may not be suitable for nonconforming feminine and masculine behaviors of LGB individuals (Gneezy, Leonard & List, 2006; Gneezy, Leonard & List, 2009). As Gneezy, Leonard and List (2006, 2009) state in their research, in some cultures, it is perfectly normal for men to stay at home and take care of their children while women go out for work. This type of culture (i.e. matriarchal culture) is not very common nowadays, however they support the claim that gender roles are shaped by cultural beliefs, customs and traditions. In other words, individuals who do not conform to traditional gender roles of a society, and feminine men and masculine women are generally stigmatized as homosexual due to the customs and traditions of that culture (e.g. patriarchal cultures) (Domanico & Crawford, 2000).

Another trigger of internalized homonegativity is religion, and specifically Christianity has been the focus of many

studies on the relation between the two (Balaji et al., 2012; Kubicek et al., 2009; Ream, 2001; Fulton, Maynard, & Gorsuch, 1999; Whitley, 2009). The relation is claimed to indicate that religious settings and faith associated behaviors with non-affirming attitudes toward homosexuality is more likely to result in homonegativity towards LGB people (Balaji et al., 2012; Fulton, Maynard, & Gorsuch, 1999; Kubicek et al., 2009; Ream, 2001). Religion is generally associated with homophobia and homonegativity on an individual level. Studies which center around the relation between religion and homosexuality give detailed explanations for the association and divided religious approaches that can be associated with anti-gay attitude, far away from acceptance and respect (Fulton, Maynard, & Gorsuch, 1999). One of them is intrinsic religious intention, which can be defined as someone drawing positive patterns about religious teachings, commitment to these teachings and living accordingly (Whitley, 2009). However, homosexual orientation is not usually welcomed by a positive attitude from intrinsic religious intentions and may cause great deal of confusion and conflict between religious and sexual identities of sexual minorities (LGB people). Ream (2001) states that people of sexual minorities carry out maladaptive and heterocentric beliefs alongside with their religious affiliations. Also, a surprising effect of intrinsic religious intentions is that LGB people seem to use their own positive religious beliefs that may arise from intrinsic intentions to overcome the stigmatization coming from non-homosexuals' religious attitudes (Ream, 2001). Since they feel overwhelmed as a result of the negative attitude, they may show a tendency to abandon their existing religious setting and focus on individual aspects of their faith (Kubicek et al., 2009). Another study, which

includes qualitative assessments and investigations, supports the claim that there is an effect of religion on sexual identity development and IH experience (Balaji et al., 2012). In this study, Balaji and colleagues (2012) disclosed some key confessions made by LGB people who support the true connection of religion and internalized homonegativity. The interviews were made anonymously in respect to participants' privacy and confidentiality.

“Even though I’m bisexual I do consider myself still being homophobic. Because when it gets to the extreme, I don’t want to be around it. Because I was raised knowing that, well...well I was raised with the belief that you know homosexuality is wrong, all this stuff. So that is still in the back of my mind (Balaji et al., 2012).”

“I come from this little small town. And everyone is always using’ those stereotypes...So it was something’ that I saw everyone else you know throwing’ rocks at. So I wanted to throw rocks at it too. So I grew up you know like a homophobe. It was like a disease to me (Balaji et al., 2012)”.

These are some anecdotal examples on how religious beliefs may turn LGB people into homophobic identities and as many studies state, this situation lead them into great deal of psychological distress (Balaji et al., 2012).

After creating a baseline about internalized homonegativity by showing evidence on its triggers; religion and ethnicity, it would be appropriate to review the effect of IH on sexual orientation disclosure and HIV-related risk behaviors. To begin with, the outness level is defined as an individual’s fitting to the stigmatizations, from fully concealing to a disclosure with nonverbal cues or verbally sharing (Balsam & Mohr, 2007; Morris, Waldo, & Rothblum, 2001). Outness rate varies from one country, culture or even one city to another; however the general rate of outness of LGB people is

not quite high worldwide. One of its main reasons is internalized homonegativity and homophobia. In their study, Balsam and Mohr (2007) concluded that psychosocial adjustment is related with outness. Wherever there is a high rate of sexual orientation disclosure, there is a lower level of internalized homonegativity in the LGB people in that environment (Herek, Gillis & Cogan, 2009; Moradi et al., 2010), and they also found that social support levels were high in participants with greater level of outness and significantly low internalized homonegativity. IH and outness levels differ between male and female LGB people. According to some studies, while fewer women had internalized homonegativity, men were more sensitive to external heterocentric stigmatization (Balsam & Mohr, 2007; Mohr & Fassinger, 2000). Outness is correlated with self-acceptance in a research (Meyer, 2003). It is stated that self-acceptance is the key to reduce internalized homonegativity and coming out to surrounding social environment.

Internalized homonegativity causes negative beliefs about one’s self and overcoming the negative beliefs with acceptance and positive evaluation has always been one of the main goals of LGB-centered therapies that help LGB people to come out. Homosexual identity formation (HIF) may come to the fore at this point. According to Herek, Cogan, Gillis & Glunt (1998) HIF, along with many other theories, has 5 stages which are a) acceptance of one’s own sexual orientation, homosexuality, b) developing positive attitude toward self, c) interaction with LGB people, d) a growing urge to disclose sexual orientation, e) mixture of homosexual identity and one’s own self-concept. Based on these stages, it can be predicted that most LGB people experience psychological stress and negative feelings

towards themselves and their homosexual experiences in general because of the failure to complete these identity formation stages (Eliason & Schope, 2007; Herek, Cogan, Gillis & Glunt 1998). The ones who could not achieve the first stage are most likely to be the ones who never step out of the closet and live their life behind the curtains. This may cause many problems in psychological, social and relational contexts, as some may get stuck in the second stage and develop negative attitude toward themselves and LGB community. People who are stuck in the second stage are most likely to harm themselves with negative feelings and become an obstacle in their own way to their sexual orientation disclosure.

It was shown that higher levels of homonegativity decrease the level of outness (Balsam & Mohr 2007; Herek, Gillis & Cogan, 2009; Moradi et al., 2010). LGB people who are not open to their surroundings about their sexual orientation and experience higher IH have higher tendency to engage in risky sexual behavior which is unprotected sexual intercourse and avoidance to look out for sexual health (e.g. safe-sex education, taking medical tests for HIV regularly) (Ross et al., 2008; Ross et al., 2011). Internalized homonegativity stems from religious settings which condemn homosexuality; therefore LGB people who are exposed to this kind of negative attitude are more likely to behave risky in their sexual life (Wilkerson, Smolenski, Brady & Rosser, 2012). Previous research about homosexual intercourse state that men who have sex with men (MSM) are more likely to search online dates for sexual partnership, much more than non-MSM community and the wide search for sexual partner increases the risky sexual behaviors (Toomey and Rothenberg, 2000). Herek and Glunt (1995) have supporting results in their study showing MSM who have higher levels of internalized homonegativity generally sees

condom-use as an offensive behavior and do not offer this to their sexual partners. Also they have found that MSM with high-level internalized homonegativity finds safe-sex practices and safe-sex education embarrassing so that they may not know how to use a condom.

It is claimed that internalized homonegativity is linked with less social support, lower self-esteem, higher symptoms of depression and unhealthy social interaction with other LGB people (Frost & Meyer, 2009) however that does not mean people with higher level of internalized homophobia are less active in sexual behaviors. Oppositely, people who are unstable in homosexual social settings do not necessarily perform unstable sexual behaviors. If it becomes anonymous, sexual interaction may occur more often than it is assumed. Unstable and confused behaviors in sex lead to high HIV risk behaviors (Huebner, Davis, Nemeroff & Aiken, 2002). A meta-analysis consisting of 16 articles conducted by Newcomb and Mustanski (2011) supports the hypothesis of positive correlation between internalized homonegativity and risky sexual behaviors. According to their study, the positive correlation between internalized homonegativity and risky sexual behaviors is stronger in the results of the publications between the years 1988 and 2008, since HIV rates and homophobia were at peak in the given time range. In addition to this information, Meyer and Dean (1995) conducted a study with 174 gay participants in the early 1990s in New York City and discovered that the relation between IH and risky sexual behaviors is much stronger in the ones who are not comfortable with their sexual orientation and who are in stress of HIV related diseases, STDs and present basic avoidance symptoms about it (Dean & Meyer, 1995).

In conclusion, it could be said that internalized homonegativity is an insidious stressor which bothers the LGB community in a hazardous way. The IH which was defined as the internalization process of society's negative beliefs, feelings and attitudes about homosexuality has significant triggers such as religion and ethnicity. Religion, in this case, may follow a dual path. Despite the fact that religion creates a background for the formation of internalized homonegativity, it also can be useful for dealing with homonegativity. Integrating one's religious beliefs and homosexuality into a new understanding may eliminate the conflict between sexual orientation and religious spirituality consequently eliminating internalized homonegativity (Rodriguez & Ouellette, 2000). In that way, theories on religion happens to be a research subject to broaden the understanding on relation between religion and internalized homonegativity. Ethnicity, on the other hand, is limited and variant in terms of validity since every ethnic origin has its own customs, traditions and views on gender. References and theories used in this article mainly focused on Latino and Black and made comparisons between these two and white people in America. For that reason, ethnicity should be subjected to in-depth consideration and should be diversified with a lot more cross-cultural research. The IH also influences important factors of homosexual life such as sexual orientation disclosure, in other words, "outness" or "coming out" and risky sexual behaviors. As it can be seen from previous literature, ethnicity and religion are two of the most powerful stressors for LGB people to develop internalized homonegativity. More research should be conducted to examine cultural and religious differences among internalized homonegativity, considering that the literature so far focuses on Black and

Latino ethnic minorities and Christians as the studied populations. The outness and sexual behaviors differ from one culture to another so further researches are also needed in order to understand the relation between internalized homonegativity and outness and risky sexual behaviors in a cross-cultural perspective.

References

- Amola, O., & Grimmert, M. A. (2015). Sexual identity, mental health, HIV risk behaviors, and internalized homophobia among black men who have sex with men. *Journal of Counseling & Development, 93*(2), 236 - 246.
- Balaji, A. B., Oster, A. M., Viall, A. H., Heffelfinger, J. D., Mena, L. A., & Toledo, C. A. (2012). Role Flexing: How Community, Religion, and Family Shape the Experiences of Young Black Men Who Have Sex with Men. *AIDS patient care and STDs, 26*(12), 730 - 737.
- Balsam, K. F., & Mohr, J. J. (2007). Adaptation to Sexual Orientation Stigma: a Comparison of Bisexual and Lesbian/Gay Adults. *Journal of Counseling Psychology, 54*(3), 306.
- Balsam, K. F., Molina, Y., Beadnell, B., Simoni, J., & Walters, K. (2011). Measuring Multiple Minority Stress: the LGBT People of Color Microaggressions Scale. *Cultural Diversity and Ethnic Minority Psychology, 17*(2), 163.
- Barnes, D. M., & Meyer, I. H. (2012). Religious Affiliation, Internalized Homophobia, and Mental Health in Lesbians, Gay Men, and Bisexuals. *American Journal of Orthopsychiatry, 82*(4), 505 - 515.

- Bridges, S. K., Selvidge, M., & Matthews, C. R. (2003). Lesbian Women of Color: Therapeutic Issues and Challenges. *Journal of Multicultural Counseling and Development, 31*(2), 113 - 130.
- Buston, K., & Hart, G. (2001). Heterosexism and Homophobia in Scottish School Sex Education: Exploring the Nature of the Problem. *Journal of Adolescence, 24*(1), 95 - 109.
- Dean, L., & Meyer, I. (1995). HIV Prevalence and Sexual Behavior in a cohort of New York City Gay Men (aged 18-24). *JAIDS Journal of Acquired Immune Deficiency Syndromes, 8*(2), 208 - 211.
- Domanico, R., & Crawford, I. (2000). Psychological Distress among HIV-Impacted African-American and Latino Males. *Journal of Prevention & Intervention in the Community, 19*(1), 55 - 78.
- Eliason, M. J., & Schope, R. (2007). Shifting Sands or Solid Foundation? Lesbian, Gay, Bisexual, and Transgender Identity Formation. *The health of sexual minorities, 3* - 26.
- Frost, D. M., & Meyer, I. H. (2009). Internalized Homophobia and Relationship Quality among Lesbians, Gay Men, and Bisexuals. *Journal of counseling psychology, 56*(1), 97.
- Fulton, A. S., Gorsuch, R. L., & Maynard, E. A. (1999). Religious Orientation, Antihomosexual Sentiment, and Fundamentalism among Christians. *Journal for the Scientific Study of Religion, 14* - 22.
- Gneezy, U., Leonard, K. L., & List, J. A. (2006). Gender Differences in Competition: the Role of Socialization. *Manuscript, University of Chicago*.
- Gneezy, U., Leonard, K. L., & List, J. A. (2009). Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society. *Econometrica, 77*(5), 1637 - 1664.
- Harris, J. I., Cook, S. W., & Kashubeck-West, S. (2008). Religious Attitudes, Internalized Homophobia, and Identity in Gay and Lesbian Adults. *Journal of Gay & Lesbian Mental Health, 12*(3), 205 - 225.
- Herek, G. H., & Glunt, E. K. (1995). Identity and Community among Gay. *AIDS, Identity, and Community: The HIV Epidemic and Lesbians and Gay Men, 2*, 55.
- Herek, G. M., Cogan, J. C., Gillis, J. R., & Glunt, E. K. (1998). Correlates of Internalized Homophobia in a Community Sample of Lesbians and Gay Men. *Journal-Gay and Lesbian Medical Association, 2*, 17 - 26.
- Herek, G. M. (2004). Beyond "Homophobia": Thinking About Sexual Prejudice and Stigma in the Twenty-First Century. *Sexuality Research & Social Policy, 1*(2), 6 - 24.
- Herek, G. M., Gillis, J. R., & Cogan, J. C. (2009). Internalized Stigma among Sexual Minority Adults: Insights from a Social Psychological Perspective. *Journal of Counseling Psychology, 56*(1), 32.
- Huebner, D. M., Davis, M. C., Nemeroff, C. J., & Aiken, L. S. (2002). The Impact of Internalized Homophobia on HIV Preventive Interventions. *American Journal of Community Psychology, 30*(3), 327 - 348.
- Kelley, T. M., & Robertson, R. A. (2008). Relational aggression and victimization in gay male relationships: the role of internalized

- homophobia. *Aggressive behavior*, 34(5), 475 - 485.
- Kubicek, K., McDavitt, B., Carpineto, J., Weiss, G., Iverson, E. F., & Kipke, M. D. (2009). "God Made Me Gay for a Reason" Young Men Who Have Sex with Men's Resiliency in Resolving Internalized Homophobia from Religious Sources. *Journal of Adolescent Research*, 24(5), 601 - 633.
- Lick, D. J., Durso, L. E., & Johnson, K. L. (2013). Minority Stress and Physical Health among Sexual Minorities. *Perspectives on Psychological Science*, 8(5), 521 - 548.
- Malebranche, D. J., Fields, E. L., Bryant, L. O., & Harper, S. R. (2009). Masculine Socialization and Sexual Risk Behaviors among Black Men Who Have Sex with Men: A Qualitative Exploration. *Men and Masculinities*, 12(1), 90 - 112.
- Malyon, A. K. (1982). Psychotherapeutic Implications of Internalized Homophobia in Gay Men. *Journal of Homosexuality*, 7(2-3), 59 - 69.
- Meyer, I. H. (1995). Minority Stress and Mental Health in Gay Men. *Journal of Health and Social Behavior*, 38 - 56.
- Meyer, H. I., & Dean, L. (1995). Patterns of Sexual Behavior and Risk Taking among Young New York City Gay Men. *AIDS Education and Prevention*.
- Meyer, I. H., & Dean, L. (1998). Internalized Homophobia, Intimacy, and Sexual Behavior among Gay and Bisexual Men. *Psychological Perspectives on Lesbian and Gay Issues*, 4, 160 - 186.
- Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. *Psychological Bulletin*, 129(5), 674.
- Mohr, J., & Fassinger, R. (2000). Measuring Dimensions of Lesbian and Gay Male Experience. *Measurement and Evaluation in Counseling and Development*, 33(2), 66 - 66.
- Moradi, B. (2010). Addressing Gender and Cultural Diversity in Body Image: Objectification Theory as a Framework for Integrating Theories and Grounding Research. *Sex Roles*, 63(1-2), 138 - 148.
- Moradi, B., Wiseman, M. C., DeBlaere, C., Goodman, M. B., Sarkees, A., Brewster, M. E., & Huang, Y. P. (2010). LGB of Color and White Individuals' Perceptions of Heterosexist Stigma, Internalized Homophobia, and Outness: Comparisons of Levels and Links. *The Counseling Psychologist*, 38(3), 397 - 424.
- Morris, J. F., Waldo, C. R., & Rothblum, E. D. (2001). A Model of Predictors and Outcomes of Outness among Lesbian and Bisexual Women. *American Journal of Orthopsychiatry*, 71(1), 61.
- Newcomb, M. E., & Mustanski, B. (2011). Moderators of the Relationship between Internalized Homophobia and Risky Sexual Behavior in Men Who Have Sex with Men: A Meta-Analysis. *Archives of Sexual Behavior*, 40(1), 189 - 199.
- Operario, D., & Fiske, S. T. (2001). Ethnic Identity Moderates Perceptions of Prejudice: Judgments of Personal versus Group Discrimination and Subtle versus Blatant Bias. *Personality and Social Psychology Bulletin*, 27(5), 550 - 561.
- Pachankis, J. E., Rendina, H. J., Restar, A., Ventuneac, A., Grov, C., & Parsons, J.

- T. (2015). A Minority Stress—Emotion Regulation Model of Sexual Compulsivity among Highly Sexually Active Gay and Bisexual Men. *Health Psychology, 34*, 829.
- Ream, G. L. (2001). Intrinsic Religion and Internalized Homophobia in Sexual-Minority Youth.
- Rodriguez, E. M., & Ouellette, S. C. (2000). Gay and Lesbian Christians: Homosexual and Religious Identity Integration in the Members and Participants of a Gay-Positive Church. *Journal for the Scientific Study of religion, 39*(3), 333 - 347.
- Ross, M. W., Rosser, B. S., Bauer, G. R., Bockting, W. O., Rugg, D. L., & Coleman, E. (2001). Drug Use, Unsafe Sexual Behavior, and Internalized Homonegativity in Men Who Have Sex With Men. *AIDS and Behavior, 5*(1), 97 - 103.
- Ross, M. W., Rosser, B. S., Neumaier, E. R., & Positive Connections Team. (2008). The Relationship of Internalized Homonegativity to Unsafe Sexual Behavior in HIV-Seropositive Men Who Have Sex With Men. *AIDS Education & Prevention, 20*(6), 547 - 557.
- Schuck, K. D., & Liddle, B. J. (2001). Religious Conflicts Experienced by Lesbian, Gay, and Bisexual Individuals. *Journal of Gay & Lesbian Psychotherapy, 5*(2), 63 - 82.
- Toomey, K. E., & Rothenberg, R. B. (2000). Sex and Cyberspace—Virtual Networks Leading to High-Risk Sex. *JAMA, 284*(4), 485 - 487.
- Wilkerson, J. M., Smolenski, D. J., Brady, S. S., & Rosser, B. S. (2012). Religiosity, Internalized Homonegativity and Outness in Christian Men Who Have Sex With Men. *Sexual and Relationship Therapy, 27*(2), 122 - 132.
- Whitley Jr, B. E. (2009). Religiosity and Attitudes toward Lesbians and Gay Men: A Meta-Analysis. *International Journal for the Psychology of Religion, 19*(1), 21 - 38.

The Relation Between Theory of Mind and Language Development: A Research Proposal to Determine the Direction of Causality

Emre Fatih

Koç University

This research proposal focuses on the relation between language development and Theory of Mind (ToM) development of children in late infancy and preschool years. Determining the direction of causality between the two cognitive abilities holds many psychological and philosophical implications, which will be addressed in the discussion section. The purpose of this research proposal is designing an experiment to ascertain this causal direction, and discussing the implications of possible results. This will be achieved by comparing two groups of preschool children subjected to ToM tests with and without explicit use of mental language. The outcome is expected to illustrate either that development of mental language precedes development of corresponding mental concepts, or that development of ToM and development of related language are simultaneous. Such a research will be a useful addition to the existing literature, which has so far ruled out the possibility of ToM development preceding corresponding language development.

Keywords: language development, theory of mind, cognitive abilities

Early childhood is a critical period for the development of language and crucial cognitive abilities, such as theory of mind (ToM) (Farrar & Maag, 2002; Wellman & Liu, 2004). ToM is the attribution of a mind, and specific mental contents (e.g., as beliefs, desires) to oneself and others, and the understanding that other people can have mental states different from one's own. The presence of both language and ToM distinguish humans from nearly all other living entities (if not all), and the development of both lies within the few years of early childhood. The strong connection between the two has been spotted and focused

upon by many researchers (Farrar & Maag, 2002; Astington & Jenkins, 1999). It has been realized that both children's own development of language and the language they have been subjected to are strong indicators of their ToM development (Grazzani & Ornaghi, 2012; Adrian et al. 2005; Ruffman, Slade & Crowe, 2002). The question that rises, after the correlation between language development and ToM development has been established, is the nature of this correlation. Is the development of ToM built upon the development of the associated language, vice versa, or are the two developing interdependently?

Many papers have indicated the correlation between language abilities and ToM performance, for many age groups and situations. ToM capabilities in these papers include attributing to one's self and also others mental states such as desires, beliefs, emotions and intentions. Grazzani and Ornaghi (2012) have conducted experiments with primary school children, examining the effect of use and comprehension of mental-state language on their false-belief (the concept that people's beliefs possibly differ from the truth) and emotion understanding. Their results indicate that comprehension of metacognitive language has strong explanatory power for children's false belief and emotion understanding, controlling for age and verbal ability, while use of such language remained relatively insignificant (Grazzani & Ornaghi, 2012). Similarly, Farrar et al. (2009) have conducted experiments with children that have Specific Language Impairment (SLI) to determine what component(s) of language development play an important role in ToM performance. Conducting such experiments with children that have SLI is particularly useful when determining the relation of specific components of language to ToM development. These specific impairments let researchers test how the said impairments affect the child's ToM performance by comparing the children with SLI to children with typically developing language abilities. This comparison opens up the possibility of investigating the effects of specific components of language on ToM development. Unlike some previous papers reporting a uniquely important role of complementation ability, such as de Villiers & de Villiers, 1999 (as cited in Farrar et al. 2009), it was observed that general grammatical and semantic abilities provide good estimates for ToM performance (Farrar

et al. 2009). This study contributed valuable information regarding typically developing children as well, for analyzing the relation between SLI and ToM performance meant having the chance to isolate specific components of language development and examining their specific effects on ToM development. This is not something that can be achieved with typically developing children.

More comprehensive longitudinal studies have been conducted concerning the relation between language development and ToM development in late infancy and preschool years by Astington & Jenkins (1999) and Farrar & Maag (2002). These studies present significant findings about the correlation between ToM and language development and the direction thereof. However, the two studies interpret their very similar findings in different ways. Astington & Jenkins (1999) cross-examine the contributions of language development and ToM development to each other in a longitudinal study. They first define the three possibilities concerning the given correlation: (1) ToM development depends on language development, (2) language development depends on ToM development, or (3) they both depend on another factor C. The third option can also be interpreted in two different ways: (1) They both depend on a common internal or external factor (Zelazo & Jacques, 1996; Dunn, 1988; Nelson, 1996 as cited in Astington & Jenkins, 1999) or (2) they are interdependent as in each contributing to the others' development via bootstrapping (Schatz, 1994 as cited in Astington & Jenkins, 1999). Astington & Jenkins (1999) conduct their experiment to determine which given account of correlation – or causation – best fits the reality. A sample of 59 children with mean age of 40.1 ($SD= 3.8$) months are tested three times for their language competence and

ToM performance in three standard ToM tests within a period of 7 months. It is expected that if language development is to lead to ToM development, language competence level at a given period would predict ToM performance at later periods. Conversely, if ToM development is to enhance linguistic development, the opposite is expected. The results indicate the first possibility (i.e. strong correlations between language levels at earlier periods and ToM scores in later periods) while the opposite is not found at significant levels (Astington & Jenkins, 1999). Astington & Jenkins (1999) infer from their results that the direction is most probably from linguistic development to ToM development, not omitting the possibility that they might still be co-dependent and language development is simply a better predictor. They have, however, ruled out the possibility of linguistic development depending on ToM development in light of their analysis. Additionally, they propose that specifically syntactic competence is instrumental for representational understanding of the mind (Astington & Jenkins, 1999). Astington & Jenkins (1999) argue that learning certain structural features of language enable the children to represent ToM related concepts in their language. An example given is the syntax of complementation, where two sentences are merged into one. The sentence "John thinks the chocolate is in the cupboard" has the sentence "the chocolate is in the cupboard" in the larger sentence "John thinks..." (Astington & Jenkins, 1999). Without the ability to use such structures, the children cannot develop and represent the concept of false-belief.

Farrar & Maag (2002) also conducted a similar experiment to assess the connection between emergence of ToM and early linguistic development. It was a within-group design. In the first period, the children were

first tested at 24 months of age for vocabulary size and then at 27 months of age for both vocabulary size and grammatical complexity. In the second period, at ages 4 and 5 ($M_{age} = 52$ Months, $SD = 4.0$ months) they were tested for their ToM scores. They were presented with one appearance-reality, one false belief, and one representational change test. What Farrar & Maag (2002) found in their results was consistent with that of Astington & Jenkins (1999). There was a strong association between language development at 2 years, and ToM scores at 4 years. More specifically, vocabulary size at 2 years of age strongly correlated with the ToM scores. Although there was also a high correlation between grammatical complexity and ToM scores, Farrar & Maag (2002) found it necessary to control this result for both vocabulary size and vocabulary growth, as there is assumed to be a strong connection between vocabulary and grammar development. However, the result of the partial correlations still suggested that grammatical development played a significant role in the acquisition of ToM, although not as strong as vocabulary size (Farrar & Maag, 2002). With the given results, Farrar & Maag (2002) pointed out that what they have reached is consistent with most of the previous work in the literature (e.g. Astington & Jenkins, 1999; Watson et al. 2001). However, unlike Astington & Jenkins (1999), Farrar & Maag (2002) opted for a different interpretation, namely that these results indicate similar social cognitive processes underlying the development of both language and ToM (Baldwin, 1993 as cited in Farrar & Maag, 2002). In other words, they argue that both depend on more general socio-cognitive development.

The above-mentioned papers will be this proposal's main reference point. The proposed research will be building on the

preliminary results these papers suggest: that either ToM and language development are co-dependent, or that ToM development relies on linguistic development. However, to understand the relation better, the effect of parental use of mental language on children's ToM development should also be investigated. Research conducted on the relation between children's ToM development and the maternal language they are subjected to, therefore provide valuable information. The amount and type of maternal language the children face can very well be an influence on their socio-cognitive abilities, which ToM and language development both depend on (Carpenter et al. as cited in Farrar & Maag, 2002).

The different data on the topic presents us with consistent results: maternal mental state utterances in general have a strong positive correlation with ToM scores of children in preschool years (Adrian et al. 2005; Howard et al. 2008; Ruffman, Slade, & Crowe, 2002). Although different types of mental state utterances had different levels of correlations (e.g. *desire* or *think* and *know* terms) with ToM scores, the most consistent correlate of ToM was the sum of all the mental state variables (Ruffman, Slade, & Crowe, 2002). Additionally, along with the frequency of such utterances, the variety thereof was also positively related with children's false belief performances (Adrian et al. 2005). Adrian et al. (2005) argue that maternal cognitive references play a crucial role in teaching the children that others are intentional agents guided by their beliefs on reality than reality itself. Interestingly, children's talk about belief and knowledge correlated more with mothers' desire talk rather than belief and knowledge talk (Ruffman, Slade, & Crowe, 2002). This indicates that children learn the concept of belief through discussion of desires. It is

inferred that frequent talk about mental states pushes children to realize the existence of subjectivity and differing perspectives (Howard et al. 2008). However, these claims are far from conclusive. Howard et al. (2008) emphasizes the need for longitudinal studies with greater age range rather than cross-sectional designs. They argue that longitudinal studies would better illustrate the progressive effect of mother and child conversations on children's mental language and ToM performance.

In the light of the given studies and remarks, I propose a longitudinal study to further narrow down the causal relationship between language development and ToM development. As stated before, it is still not determined whether ToM builds upon linguistic development or the two are co-dependent/interdependent, although a third possibility is arguably ruled out (Astington & Jenkins, 1999; Farrar & Maag, 2002). There is still not enough data to choose between the interpretations of Astington & Jenkins, and Farrar & Maag. This research might give the needed preliminary data to weigh the two interpretations. What I want to do is to distinguish the knowledge of mental concepts from knowledge of mental language, if that is possible at all. It is conceivable that a child might have affinity with a word or a general structure, knowing when and how to use it, although not grasping it's meaning. This may very well be the case with mental language and mental concepts. The child for example might be attributing such language to certain behavioral situations rather than mental states. If this is possible, it may imply that children first learn the mental language and then construe the concepts upon the language. If however, the mental language and mental concepts are indeed inseparable, this will then imply the development of both simultaneously. What I want to test is whether

ToM tasks can be successfully performed; only being familiar with the situation and the language, although not having understood the concepts the task is actually testing for. Such a test will have two uses in the literature. First, the ToM tasks themselves will be tested for validity with respect to language competence. In other words, these tasks will be tested to see whether they measure ToM performance, or ToM related language competence. The current literature seems insufficient at addressing this possibility of invalidity from this perspective. The second and more substantial use will be that such a test will provide valuable insight to the developmental relation between ToM and language.

The proposed experiment aims to test for this by applying the ToM tasks designed by Wellman & Liu (2004). Wellman & Liu started with the assumption that typically developing children follow a predictable trajectory in their ToM development. With this assumption, they had made a large-scale meta-analysis to order the various ToM tasks according to their relative difficulty. In this case, the more difficult the task, the older the children would have to be (i.e. have more developed ToM, to pass that task). The meta-analysis yielded a consistent sequence of ToM tasks (Wellman & Liu, 2004). The findings were as follows:

“Young children can first (a) correctly judge that two persons have different beliefs, and (b) correctly judge how a person’s action follows from their beliefs (in contrast to the child’s own opposite belief). Only later can children correctly make the same judgments when they do know what is true and hence can (c) correctly judge that one person’s belief is true and the other person’s belief is decidedly false, and (d) correctly judge how a person’s actions mistakenly follow from a false belief.”

To test their findings in the meta-analysis,

Wellman & Liu (2004) then conducted an experiment of their own where groups of 3-, 4-, and 5-year olds were given seven different ToM tests. These were tasks on (1) diverse desires, (2) diverse beliefs, (3) knowledge access, (4) contents false belief, (5) explicit false belief, (6) belief emotion, and (7) real-apparent emotion respectively. Consistent with the assumptions and the results of the meta-analysis, the tasks were found to align from easy to difficult in the above given order, diverse desires task being the easiest, and real-apparent emotion task the hardest (Wellman & Liu, 2004). The majority of children that were able to correctly answer a certain task were also able to complete the tasks before it, and conversely if a child failed a certain task, they would fail the ones coming after it. That is, if a child is successful at task 5 for example, it is expected that they would be successful with previous tasks (1-4), and if the child is unsuccessful at task 3, it is with high probability that they also fail at upcoming tasks (4-7). Also, as expected, the older children were successful in more tasks than the younger ones (Wellman & Liu). Hence, this study presents us with a useful tool for our research.

Method

Participants

Two groups of 3-year old children should be recruited as experimental groups, to be tested annually (once they are 3, 4, and 5 years of age). General verbal competency should be measured to control the results for differing linguistic abilities. Additionally, two control groups of primary school children should be recruited to determine the possible difference of difficulty inherent to the tasks given to first two groups. This way, this difference can also be controlled for.

Procedure & Statistical Analysis

To distinguish the effect of mental language on ToM scores, two different

versions of the same ToM tasks should be presented to two different experimental groups. One group (g1) should listen to the ToM tasks with explicit usage of mental language, while the other group (g2) listens to the same ToM tasks with minimal use of mental language. This way, any positive effect of ‘knowing’ the language can be differentiated. To see what should be taken as mental language, the study of Ruffman, Slade, and Crowe (2002) can be used as an example. If the development of mental language does not precede the development of corresponding mental concepts, g1, which hears the ToM tasks with mental language, should have no advantage over g2, and therefore score similarly. If however, the mental language develops prior to the mental concepts, hearing the same ToM task with explicit mental language should have a positive effect on scores, thus resulting in g1 scores being significantly higher than that of g2. The experiment hence should be conducted as a one-tailed test. The null-hypothesis suggests that there is no significant difference in scores. The null-hypothesis is rejected if g1 scores significantly higher than g2.

In order to make the study more thorough, the suggested experiment should be done in three periods (each period being a year long). In each period all seven of the scaled ToM tests should be administered. Making a longitudinal study and administering all seven tasks have valuable additions to the analysis. By making a longitudinal study, we can observe how the scores change as the children develop further. Assuming that null-hypothesis is false, the expectancy is that moving through the periods, the scores of the two groups should converge for any single task. Convergence is expected because as the children age, their performance on a specific component of ToM will gradually increase, thereby decreasing

the relative advantage of being exposed to oral mental language for g1. In the end, the understanding of the mental concepts should catch up with the understanding of the mental language. In addition, giving all seven tasks allows us to observe the precedence of the mental language over the mental concepts. If the null-hypothesis is to be rejected, we should expect to see a divergence in the scores of the two groups as the difficulty of the task increases, at any given time. With an easier task, an earlier developed ToM content is tested, which means there would be more time for the concept to build upon the language. Harder task means a more recently developing ToM component, where the concept has not yet been built on the language. Therefore, differences in scores should increase as harder tasks are given. It should also be checked, whether the results from different tasks are consistent.

The results should be coded once for the effect of age (on x-axis) on ToM score (on y-axis) for each task as two separate functions for g1 and g2. This will illustrate the expected convergence over increasing age between the two groups. It should also be coded for the relation between difficulty of ToM tasks and the score difference. X-axis will have the ToM tasks lined ascending in difficulty, while y-axis will again code ToM scores of g1 and g2 for each ToM task at a specific given time (t). This will present the expected divergence between the scores of g1 and g2 as the difficulty of task increases.

Discussion

Determining the direction of correlation between language development and ToM development is a critical step both philosophically and psychologically for two reasons. First, the result can be interpreted within its own scope, that is, we can examine

what the results of the study would mean for ToM development, and our understanding of intentional agents. Secondly, any conclusion we reach can be interpreted under a broader perspective. We can look at how any conclusion we reach about ToM development and language development relation can be applied to the discussion about the relation between developments of the mind in general and of language. Neither discussion has more priority than the other, because as their scopes differ, their areas of effect also differ.

One specific, which can be discussed, is the direction of causality between ToM and language. As mentioned, there are three conceivable directions of causation: unidirectional causality either from language to ToM or vice versa, or bidirectional causality between the two.

One assumption can be that mental language development depends on ToM development. This would indicate that independent from any language development, ToM develops under the right social environment. Hence, independent from any linguistic effects, there is a universal process for development of ToM in the mind of children. In other words, knowing of other minds is not necessarily a socially learned skill, but may well be a fundamental aspect of the brain. Mental language only develops because the child developing this fundamental component of the mind requires that language to address the rising concepts. Cognitive sciences have been working to locate which part of the brain correlates with ToM development, and try to understand how ToM can be encoded in the brain with papers such as Apperly et al. (2005) and Callejas et al. (2011). This literature can be expanded with further neurological research focusing more on the relation between areas of language and ToM. At the same time philosophy would be striving to answer the question: how can one

mind by simple biological development come to know other minds and how does the conceptual development of mental language can be explained.

If we assume the opposite, that ToM development depends on language development; we have a very different situation at hand. If we claim that, we mean that ToM builds upon the given lexical and syntactic structures that a child possesses. Development of mental language becomes a necessary condition for ToM. Not only that, but having different lexical and syntactic structures (as in different languages) might strongly influence the way ToM develops, which implies many different understandings of subjectivity, mind, and intentional states by many different agents. It might be, that how the ToM develops strongly depends on the variance in both lexical and structural differences. The society and culture we are in influence our language, which in turn would strongly limit and shape our understanding of other minds. Psychology in this case must study further whether and how much differing linguistic structures affect ToM development. Initial results have been somewhat significant concerning ToM development with respect to linguistic and socio-economic differences, but more research is clearly needed (Cutting & Dunn, 1999; Liu et al. 2008). It must also investigate how much ToM's of different people can diverge because of these differences. On the other hand, philosophy will have to investigate the scope of limitations caused by linguistic boundaries. Can we, in spite of the linguistic constraints, find a common understanding of ToM? Or do the limitations of the given schema force us into subjective understandings of subjectivity at best? A unidirectional causality from language to ToM therefore represents a very strong interpretation of social constructivism, where the understanding of other minds is

only definable under a given language and is inconceivable without one.

What if there was bidirectional causality (or correlation)? This would negate strong social constructivism while still keeping the dependence of ToM on mental language. The inseparability of ToM development from language development still indicates a social constructivist view, although weaker. ToM development still would be through lingual exchanges in social environments. However, the simultaneous development supports the view that language – and the social environment – is not a unidirectional influence. While mental language development is necessary for mental concept development, it also requires the development of mental concepts themselves. This implies a more objective, physiological rather than social component in ToM development. This time however, studying the mental concepts or mental language development independent from the other becomes problematic. A problem of reflexivity arises: as mental concepts hardly become something over and above their corresponding language, it becomes near necessity to investigate ToM within a linguistic construct. The problem would be of separating the capacity of ToM from the vocabulary and structure used to address it.

As can be seen, no matter which interpretation is favored, additional, even more complex questions rise as follow-ups. When we direct our focus to the even broader topic of language development and development of mind, the questions double in complexity and number. The implications of strong social constructivism become even bolder and broader. The whole system that we call the mind can be argued to be trapped in the given language, time, and social structure, never able to break through them to reach an objective understanding. The interdependency

between the developments of the two and their inseparability become even more complicated. If it is inseparable from language, can we really conceive of a mind without conceiving of its language? If we can't what does this imply about the nature of the mind and of language itself? If, on the other hand, the mind develops ahead of language, we can expect similar patterns of development in all humans, irrespective of language variations. How and why then, are there language variations at all?

Whatever answers and questions there may be, the proposed research should still be useful offering more detailed data on the correlation between ToM development and mental language development, and more specifically focusing on the direction of the correlation. This is an important point to focus on, because only when we start to favor one interpretation more than the other and find consistent results for that interpretation, can we further tackle these even more complex, and perhaps exciting, questions regarding mind and language.

References

- Apperly, I. A., Samson, D., & Humphreys, G. W. (2005). Domain-specificity and theory of mind: evaluating neuropsychological evidence. *Trends in Cognitive Sciences*, 9(12), 572-577.
- Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, 35(5), 1311-1320.
- Callejas, A., Shulman, G. L., & Corbetta, M. (2011). False Belief vs. False Photographs: A Test of Theory

of Mind or Working Memory? *Frontiers in Psychology*, 2.

- Cutting, A. L., & Dunn, J. (1999). Theory of Mind, Emotion Understanding, Language, and Family Background: Individual Differences and Interrelations. *Child Development*, 70(4), 853-865.
- Farrar, M. J., Johnson, B., Tompkins, V., Easters, M., Zilisi-Medus, A., & Benigno, J. P. (2009). Language and theory of mind in preschool children with specific language impairment. *J Commun Disord*, 42(6), 428-441.
- Farrar, M.J., Maag L. (2002). Early language development and the emergence of a theory of mind. *First Language*, 22, 197-213.
- Grazzani, I., & Ornaghi, V. (2012). How do use and comprehension of mental-state language relate to theory of mind in middle childhood? *Cognitive Development*, 27(2), 99-111.
- Howard, A. A., Mayeux, L., & Naigles, L. R. (2008). Conversational correlates of childrens acquisition of mental verbs and a theory of mind. *First Language*, 28(4), 375-402.
- Juan E. Adrian, Rosa A. Clemente, Lidon Villanueva and Carolien Rieffe (2005). Parent-child picture-book reading, mothers' mental state language and children's theory of mind. *Journal of Child Language*, 32, pp 673-686
- Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, 44(2), 523-531.
- Ruffman, T., Slade, L., & Crowe, E. (2002). The Relation between Childrens and Mothers' Mental State Language and Theory-of-Mind Understanding. *Child Development*, 73(3), 734-751.
- Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development*, 75(2), 523-541.