Summary
A Study on the Reliability and Validity of the Turkish Free Association Norms

Ümit Akırmak

İstanbul Bilgi University

Mehmet Altan Orhon

İstanbul Bilgi University

Free association helps us in mapping the associative structure of words in semantic memory by indexing the strong associations of a given word (Deese, 1965; Nelson, McEvoy, & Schreiber, 2004). Understanding the associative relations among the words is crucial especially for studies that make use of words as stimuli since the characteristics of the words have been shown to directly influence the results of the experiments (Nelson, Fisher, & Akirmak, 2007; Nelson, Schreiber, & Xu, 1999). For this reason, pre-existing associative knowledge about the words is usually utilized to control various word attributes across experimental conditions. In addition to controlling for experimental conditions, associative knowledge is also used to predict memory performance in various tasks (Deese, 1965; Nelson, Fisher, & Akirmak, 2007), false memory rates in the Deese–Roediger–McDermott (DRM) paradigm (McEvoy, Nelson, & Komatsu, 1999), as well as differences in behaviors and attitudes of different populations (Reich & Goldman, 2005; Stacy, 1997). As it was shown, the free association task has broad applicability in psychological research and is still being widely used in different types of research.

In the free association task, participants are provided with target words and are asked to produce the first word that comes to their mind in association with each target word. A prior study by Tekcan and Goz (2005) collected free association norms for 600 Turkish words; however, the reliability and validity of the responses they obtained have not been tested by an additional study. When the studies on free association norms are examined in the literature (e.g., De Deyne, Navarro, & Storms, 2013; Nelson, McEvoy, & Schreiber, 2004), it can be seen that norming procedures often involve reliability and validity analyses through new studies. The purpose of the present study was to examine the reliability and validity of the existing free association norms for the Turkish words (Tekcan & Goz, 2005). Two studies were conducted to this end as part of the present study.

These studies focused on associative information, which was indexed by free association, such as associative strength, set size, and heterogeneity. The present study first reviewed these concepts and theoretical models that were specifically aimed to explain free association and cued recall tasks. The relative strength hypothesis (Maki, 2008; Nelson, McEvoy, & Dennis, 2000) for predicting response probabilities in a free association task and "Processing Implicit and Explicit Representations II" (PIER II; Nelson, McKinney, Gee, & Janczura, 1998) for predicting memory performance in cued recall task were explained in detail. The specific information on studies as well as the findings were presented.

Study 1

The first study aimed to assess test-retest and internal reliability of the associative structure of the Turkish words as measured by free association. For this purpose, 100 words were randomly chosen from the existing free association normative database (Tekcan & Goz, 2005) and were used as materials in the present study. The standard free association task was applied to 105 undergraduate students. In addition to the standard method, an online version of the task was designed and was applied to 59 undergraduate students through an online survey site namely *Survey Monkey*. The analyses were mainly focused on two variables of interest: associative strength and associative set size. To assess the reliability of the existing association norms, the newly obtained values of strength and set size through standard and online methods were compared with the existing values. In addition, as it was suggested by De Deyne et al. (2013), heterogeneity values for the three databases were computed and compared as an alternative index of reliability. The results indicated that for all three databases, the associative strength values decreased consistently with the rank of a given associate. In addition, the first associates showed a high test-retest

Address for Correspondence: Dr. Ümit Akırmak, İstanbul Bilgi University, Faculty of Social Sciences and Humanities, Department of Psychology, Eski Silahtarağa Elektrik Santralı Kazım Karabekir Cad. No:2/13 Eyüp / İstanbul
E-mail: umit.akirmak@bilgi.edu.tr

correlation of associative strength (r = .72) and high internal consistency as shown by comparison of the standard task with the online task (r = .89). Finally, the three databases also showed a strong positive correlation of heterogeneity values, indicating that they produced similarly different associative responses for the target words. Thus, these findings implied that free association reliably indexes associative word knowledge. More importantly, existing Turkish free association norms were found to be reliable as indicated by similar pattern of responses obtained from a different sample, measured at a different time and by utilizing a different method.

Study 2

The purpose of the second study was to examine whether the existing normative free association database could be used to predict memory performance in a new task, based on the expectations from similar studies in the literature. Specifically, according to PIER II, cues that have stronger forward connections to their targets should result in better correct cued recall as compared to cues that have weak or no forward connections when all other variables are controlled (Nelson et al., 2007). The same advantage is also present for the cues that produce fewer associates in free association as compared to the cues that produce many associates (Nelson et al., 1998). These predictions were tested in a cued recall experiment designed to vary some of the word characteristics that were measured by free association. In the cued recall task, participants first studied a list of words (targets) and then were asked to recall these words with the help of associatively related words (cues) that were not shown during the experiment. Associative strength was manipulated between subjects and cue set size was manipulated within subjects in this cued recall experiment. For this purpose, 48 cue words and their corresponding targets were chosen from the normative database (Tekcan & Goz, 2005) and were allocated into four different word lists. There were two word lists in each between-subjects condition in order to examine the generalizability of the findings across chosen materials. Sixty-four undergraduates participated in the study, in exchange for course credit. Because of the structure and the extent of the existing word norms, only various word attributes for the cue words could be controlled in the present study. These attributes included word frequency, concreteness, and imagery. The results of the experiment indicated that stronger associative strength between the cue–target pair resulted in higher probability of cued recall as compared to weaker associative strength between the cue–target pair. This result replicated earlier findings on associative strength and added to the generalizability of PIER II by confirming

its predictions with material from a different language's free association norms. However, the manipulation of the cue set size did not result in the expected advantage for the cues having a smaller associative set size. A close inspection of the findings revealed that one set of lists confirmed PIER II's predictions on cue set size effects while the other set contradicted it, implying that there might be a confound in the construction of the study lists. Further analyses revealed that this effect was not due to confound variables controlled within the study. Although this result seems to contradict PIER II's predictions in general, the apparent contradiction can be explained by uncontrolled word attributes belonging to the target words. Specifically, according to PIER II, cue set size effects result from an intersection of the cue and target's associative structures. Thus, in order to fully predict the cue set size effects, knowledge regarding the target's associative structure must be present; this was not the case for the present study. Overall, the findings implied that Turkish word association norms possess predictive validity, especially when associative strength was considered. However, more research on associative set size is needed to be able to entangle how the associative structure of the target influences correct recall. In order to be able to assess the contribution of the associative structure of the target words on cued recall and to map the associative structure of these words, it is needed to conduct additional norming studies on the responses of the normed words in the normative database.

General Discussion

The purpose of the present study was to examine the reliability and validity of Turkish word association norms. Study 1 evaluated the test-retest reliability of the three databases (existing norms, newly collected norms through standard method, and newly collected norms through an online survey) by comparing associative strength, set size, and heterogeneity values. The findings of Study 1 indicated high test-retest reliability. Although there exist differences in the particular associates produced, the overall pattern of associates provided in response to a target word was highly similar. The present findings showed for the first time that Turkish word associations (Tekcan & Goz, 2005) are reliable despite the time elapsed and different population utilized. The validity of the normative database was examined by designing a cued recall experiment based on the predictions of a mathematical model of cued recall task, namely PIER II (Nelson et al., 1998). The results of Study 2 partially supported the validity of the existing norms by showing that associative strength reliably predicted cued recall memory performance. The results in regard to the cue

set size failed to provide full support mainly because of the insufficient knowledge for the associative structure of the target words utilized in the second study. In other words, cue set size effects are based on the intersection of cue and target's associates, and thus the knowledge of both cue and target and their associative structures are needed in order to reliably demonstrate their effects. In the absence of such information, it is not possible to accurately predict the recall advantage for the cues with smaller associative sets as compared to cues with larger associative sets. The findings obtained from one set of lists confirmed the cue set size effects, while the results obtained from another set of lists failed to provide a support. Overall, the results of the cued recall study was promising in that it confirmed the predictive validity of the associative strength values indexed by free association norms collected about a decade ago. In addition, present findings supported the cross-cultural validity of PIER II in predicting cued recall performance based on the word attributes measured by free association norms from a different language.

Word association norms are used in various tasks and for various purposes. We believe collecting further word association norms will be a productive endeavor as long as the norms serve as reliable predictors of task performance. As an example, the seemingly contradictory results of Study 2 might be derived from the insufficient knowledge on word associations. Thus, extending the normative database for word associations may enable researchers, who are working with words as stimuli, to apply experimental control more reliably in their studies. In addition, an extended database for word associations may enable researchers to make use of different research methodologies and designs, which, currently, can not be utilized without the potential confounds mentioned in the present study as lack of associative knowledge.