Learning and Individual Differences xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Learning and Individual Differences

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

Measuring two aspects of emotion recognition ability: Accuracy vs. sensitivity ☆

Dmitry Lyusin [a,b,*], Victoria Ovsyannikova [a]

[a] National Research University Higher School of Economics, Myasnitskaya 20, Moscow 101000, Russia
[b] Russian Academy of Sciences, Institute of Psychology, Yaroslavskaya 13, Moscow 129366, Russia

ARTICLE INFO

Article history:
Received 31 January 2014
Received in revised form 15 January 2015
Accepted 29 April 2015
Available online xxxx

Keywords:
Emotion recognition
Emotion intelligence
Videotest of Emotion Recognition

ABSTRACT

The present paper aims at showing the necessity to distinguish two aspects of emotion recognition ability, accuracy of the recognition of emotion types that constitute the emotional state of the perceived person and sensitivity to the intensity of the perceived person's emotions. A new technique that measures these two aspects of emotion recognition, the Videotest of Emotion Recognition, is proposed. The accuracy and sensitivity indices provided by the Videotest of Emotion Recognition have high reliability and yield different correlation patterns with other cognitive and personality variables.

1. Introduction

Emotion recognition has been widely studied for decades in psychology. In modern psychology, emotion recognition is often conceptualized and measured in the frame of emotional intelligence research. Broadly, emotional intelligence refers to the set of abilities that allows understanding and managing of emotions. Emotion recognition is widely considered to be one of the basic emotional intelligence components. The well-known emotion intelligence model proposed by Mayer, Salovey, Caruso, and Sitarenios (2001) identifies four branches of emotional intelligence; two of them, Emotion Perception and Emotion Understanding, are related to emotion recognition. Emotion Perception includes skills concerned with accurate detection and identification of emotions in oneself and others. Emotion Understanding concerns the ability to understand relationships between emotions, emotion language and signals conveyed by emotions. According to this model, four branches are ordered hierarchically, the basic branch being Emotion Perception (Salovey & Grewal, 2005). It seems that distinguishing Emotion Understanding from Emotion Perception is artificial and has an intuitive rather than theoretical background.

Another approach to the conceptualization of emotion skills, proposed by Scherer and Scherer (2011), understands emotion perception as one of the three major domains of emotional competence along with emotion production and emotion regulation. Emotion perception is considered a central socio-emotional competence essential for many different types of occupation.

One of the important directions in emotion recognition research is developing methods for measuring emotion recognition ability. Most of these methods focus on accuracy of emotion recognition. The present paper aims at showing the necessity to distinguish between the two aspects of the ability to recognize emotion, namely accuracy and sensitivity; a technique for measuring accuracy and sensitivity is also proposed.

1.1. Tests for measuring emotion recognition ability: diversity and problems

The number of studies on measuring emotion recognition ability has been growing in the recent decades. Most of the new measurement instruments have been developed in the context of emotion intelligence assessment. Two types of assessment methods are traditionally distinguished in the research on emotional intelligence, objective tests and self-report questionnaires. They correspond to the two types of emotional intelligence models that are usually called ability and mixed models (Mayer, Salovey, & Caruso, 2000). Ability models understand emotional intelligence as a set of cognitive abilities and competencies analogous to other types of intelligence such as verbal or spatial. Mixed models, also called trait models, define emotional intelligence more broadly, as an array of cognitive, personality, and motivational traits that provide better emotion understanding and management, and finally result in higher levels of adaptation and well-being of an individual. For measuring emotional intelligence, proponents of ability models use objective tests similar to traditional intelligence tests with answers that can be assessed

☆ This study was implemented in the framework of the Basic Research Program at the National Research University Higher School of Economics in 2015.
* Corresponding author at: Leningradskoe shosse 48-1-29, Moscow 125212, Russia.
E-mail addresses: ooch@mail.ru (D. Lyusin), v.ovsyannikova@gmail.com (V. Ovsyannikova).

http://dx.doi.org/10.1016/j.lindif.2015.04.010
1041-6080/

as right or wrong. Mixed models proponents prefer self-report questionnaires similar to personality inventories. Some exceptions from this correspondence between the two types of models and approaches to measurement are possible. For example, the EmIn Questionnaire developed by Lyusin (2006a, 2006b) that will be described below is based on the ability model. The author claims that it measures perceived emotional intelligence, understood as a cognitive ability, rather than personality traits.

The limitations of self-report assessment are broadly known; hence this paper will focus on objective tests that evaluate emotion recognition ability independent of an individual's self-concept and beliefs about his or her behavior. There is a large diversity of such tests in modern psychology. They differ in stimuli, item formats, indices, and scoring procedures. For instance, stimuli can be photographs of facial expressions, videos with various types of behavior, voice recordings, vignettes describing emotional situations, and even thoroughly non-human stimuli such as geometric figures.

The problem of scoring is one of the hardest in performance-based assessment of socio-emotional abilities. Unlike traditional intelligence tests, there are no obvious logical foundations for establishing correct answers in most emotion recognition tests. Three major approaches to scoring have been suggested, namely expert, consensus, and target scoring. Expert scoring is based on expert opinions about the correct or best choice among suggested answers. The main difficulty is to decide who has expertise in this domain. In most cases, emotion researchers are suggested for this role, but it is often questioned if they or any other professionals such as psychotherapists, counselors, and actors qualify as emotion experts. Some authors even claim that the emotion domain is one of those ill-defined knowledge domains where no objective standards for verification exist and, therefore, no qualified experts can be suggested (Legree, Psotka, Tremble, & Bourne, 2005). Consensus scoring is based on the opinion of the majority of the participants about correct answers. It is often supposed that consensus scoring reflects cultural biases in beliefs about emotions. Moreover, it is regarded as logically unacceptable to establish correct answers to the intelligence test items, especially to the difficult ones, on the basis of the consensus opinion. In target scoring, the correct response is set by a target person who creates the stimuli. These target persons can be actors portraying emotions for photographs or voice recordings, authors of the vignettes who define a priori which emotion should be experienced by a certain character, etc. Target scoring can be applied only to a limited range of stimuli, and it can always be questioned if the target emotion was adequately portrayed or expressed in the stimuli. All three approaches have their own limitations, but they are used in psychological research and assessment for the lack of better solutions.

An important feature of emotion recognition items, as well as of any emotional and social abilities items, is the difficulty in establishing one correct response. Several responses to the same item can often be regarded as correct with different levels of confidence. This situation is quite normal for the psychological content being measured since emotional states are often ambiguous and constitute a mixture of various emotion types. The stimuli cannot represent all individual and situational features that result in a certain emotional state. Two important consequences result from this. First, it makes sense to use rate-the-extent format of responses similar to the Likert-type scales, rather than just to classify responses as correct and incorrect. Secondly, the unidimensional format of responses when a participant estimates the presence of only one emotion in the stimulus is less appropriate as compared to the multidimensional format that allows estimating the presence of an array of emotions in the stimulus.

Different approaches to scoring and different response formats (unidimensional or multidimensional) are used in modern emotion abilities tests. The following brief review of emotion recognition tests summarizes the main tendencies in this field.

One of the most prominent early techniques for emotion recognition is the Profile of Nonverbal Sensitivity (PONS; Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979). It consists of twenty audio/video recordings in which one female person represents twenty attitudes (such as expressing jealousy, asking for a favor). The participant must assess the attitude expressed by the character. Attitudes are set initially by the test developer and are classified as dominant versus submissive and positive versus negative. Each recording is represented by eleven channels of expression (face, speech, etc.). The 220 portrayals are presented to the participant in a fixed order. For each portrayal, the participant is required to select one of two alternative answers. The accuracy index is calculated as the percentage of correct answers of the total number of test stimuli.

The Diagnostic Analysis of Nonverbal Accuracy was designed to assess the sensitivity to nonverbal expressions of emotions (DANVA; Nowicki & Duke, 1994). Twenty-four photographs of facial expressions and 24 voice recordings of four emotions (anger, fear, joy, sadness) are used as stimuli. Each emotion category is presented in two intensities, low and high. The emotions are portrayed by professional actors. The participant has to choose one of the four emotion categories for each stimulus. The accuracy scores are calculated as the percentage of correct responses separately for both types of stimuli and for the whole test.

A notable feature of the Japanese and Caucasian Brief Affect Recognition Test is the use of the images of people of different races as stimuli (JACBART; Matsumoto et al., 2000). The photographs of European and Japanese facial expressions of seven basic emotions (anger, joy, sadness, contempt, disgust, fear, and surprise) are presented to participants who have to assess the presence of each of the seven emotions in the portrayals by means of nine-point scales. The average values for each emotion category obtained in the American sample are considered to be standard. Accuracy scores are calculated as correlations between the participants' responses and the standard estimates. An interesting feature of the technique is the possibility to calculate different accuracy scores separately for each emotion category, for different races and sexes.

The most famous measure of emotion recognition is the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT; Mayer, Salovey, Caruso, & Sitarenios, 2003). The test is based on Salovey and Mayer's model of emotional intelligence that regards it as a set of hierarchically organized cognitive abilities. The MSCEIT consists of four subtests. The first and third subtests, Emotion Perception and Emotion Understanding, measure abilities related to emotion recognition. The Emotion Perception subtest includes two types of tasks with photographs of facial expression and pictures of landscapes and abstract designs as stimuli. The participant must assess the degree of presence of several emotions in each stimulus using Likert five-point scales. The Emotion Understanding subtest consists of the Blends task and the Changes task. In the Blends tasks, the participant must identify which emotions will result from the blend of several other emotions and select one of the response options. In the Changes tasks, the participant must select the emotion from the list of emotions that may result from the situation described. The weights based on expert and consensus ratings are attributed to each response option. An accuracy index is calculated by averaging the weights of the responses selected by the participant.

Recently, the Emotional Intelligence Measure (AEIM; Warwick, Nettelbeck, & Ward, 2010) was developed, which is, actually, a revised version of the MSCEIT. The two scales, Emotion Perception and Emotion Management, have been changed. The principles of stimuli selection and scoring methods are similar to the MSCEIT.

The Situational Test of Emotional Understanding (STEU; MacCann & Roberts, 2008) consists of the descriptions of situations related to different emotions. The STEU items were developed according to Roseman's appraisal theory of emotions (Roseman, 2001). The test authors set the correct responses on the basis of this theory. The accuracy index is calculated as the percentage of correct responses.

The Multimodal Emotion Recognition Test (MERT; Bänziger, Grandjean, & Scherer, 2009) consists of the presentations of expressions of five emotion families in four formats, video with sound, video without sound, audio without image, and photo taken from video. Emotional

Please cite this article as: Lyusin, D., & Ovsyannikova, V., Measuring two aspects of emotion recognition ability: Accuracy vs. sensitivity, *Learning and Individual Differences* (2015), http://dx.doi.org/10.1016/j.lindif.2015.04.010