Construct Validity of the Emotion Matching Task: Preliminary Evidence for Convergent and Criterion Validity of a New Emotion Knowledge Measure for Young Children

Judith K. Morgan, Carroll E. Izard and Kristen A. King, University of Delaware

Abstract

Current emotion knowledge (EK) measures examine only one component of the multifaceted construct. We examined the reliability and the construct validity of a new measure of EK, the emotion matching task (EMT). The EMT consists of four parts which measure the components of receptive EK, expressive EK, emotion situation knowledge, and emotion expression matching. First, we compared the EMT and its parts to two widely used EK measures—the Kusche emotional inventory (KEI) and Denham’s affective knowledge test (AKT, 1986) in order to establish convergent validity. The EMT and its four parts were strongly correlated with both measures of EK. Regression analyses revealed moderate to strong predictive validity for EMT. Compared with KEI and AKT, the EMT was a more robust predictor of teacher-rated emotion regulation and parent reported effortful control. Compared with KEI and AKT, the EMT correlated similarly with verbal ability and age.

Keywords: emotional knowledge; construct validity

Comprehensive measurement of emotion knowledge (EK) is needed in order to assess skills that help determine peer relations and social competence. Currently, there are several measures of EK, but they do not measure some of the key facets of the construct. Because of the complexity of this construct, it is important to develop and validate measures of EK that tap into several of its multiple components. Moreover, past and current EK measures are often highly correlated with age and verbal ability, and may be confounded by these constructs. Consequently, new EK measures that are more appropriate for younger children or those with delayed language-related cognitive ability are needed. The current study evaluates the validity of a new measure of EK for young children, the emotion matching task (EMT; Izard, Haskins, Schultz, Trentacosta, & King, 2003), that attempts to address these issues.

Correspondence should be addressed to Judith K. Morgan, Department of Psychology, University of Delaware, Newark, DE 19716, USA. Tel: 302-831-2700; Email: jkmorgan@psych.udel.edu

© Blackwell Publishing Ltd. 2009. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.
**The Construct of EK**

EK includes the ability to recognize expressions of emotions, label expressions of emotions and understand the causes and consequences of emotions. These abilities are often termed receptive EK (recognizing emotion expressions), expressive EK (labeling expressions), and emotion situation knowledge (understanding emotion as a causal process), respectively (Ackerman & Izard, 2004; Denham, 2003; Fine, Izard, Mostow, Trentacosta, & Ackerman, 2003; Schultz, Izard, Ackerman, & Youngstrom, 2001). Although understanding emotions may involve other components, we believe these abilities to be most important in enabling effective communication and adaptive use of emotion expression in young children (Izard, Trentacosta, King, Morgan, & Diaz, 2007). EK has been related significantly to prosocial behavior, higher social competence, and emotion regulation in 3- to 7-year-old children (Denham et al., 2003; Izard et al., 2001; Mostow, Izard, Fine, & Trentacosta, 2002; Shields, Dickstein, et al., 2001). Higher EK may be related to effortful control and higher academic competence in young children (Fabes, Martin, Hanish, Anders, & Madden-Derdich, 2003; Trentacosta, Mostow, Izard, & Fine, 2007).

**Components of EK**

The construct of EK consists of several components. Those of concern here are receptive EK, expressive EK, and emotion situation knowledge. Receptive EK usually emerges first in early development. It first manifests itself in the ability to recognize emotion signals in the facial expressions or vocal tone of others (Denham, 2003). Early on, infants reference human faces to learn and share emotion expressions. These expressions can communicate feelings of joy, sadness, anger, and fear to the young infant (Izard et al., 1995; Trevarthen, 2005). As expected in this period of development, some researchers have found children’s ability to recognize emotions to be unrelated to their verbal ability (Dunn, Brown, & Beardsall, 1991).

Expressive EK develops fairly rapidly in the preschool years but can start as early as 18 months of age (Denham, 2003; Odom & Lemond, 1972). Toddlers may witness their parents expressing emotion feeling and using emotion labels in conjunction with these emotion expressions. As a result, toddlers may learn to attach these labels to facial expressions, leading to the development of expressive EK (Izard, 1991; Schultz et al., 2001). Cognitive ability, particularly verbal skill, plays a role in expressive EK, such that greater verbal skill aids in greater ability to label emotion expressions accurately.

Emotion situation knowledge, understanding which situations tend to elicit which emotions (e.g., happiness, sadness, anger, and fear) emerges during the preschool years and continues to develop through the grade school years (Brody & Harrison, 1987; Denham, 2003; Denham & Couchoud, 1990; Stifter & Fox, 1987). However, children as young as two may talk about their own and others’ emotions as well as the causes and consequences of those emotions in casual conversation (Dunn, 1994). Such emotion discourse contributes to the early stages of the development of emotion situation knowledge.

There may be different cultural values of the importance of emotion discourse in the home. Wang (2003) found 3- to 6-year-old American children to have higher emotion situation knowledge than Chinese children of the same age due to less discourse about emotions between Chinese parents and their children. Low-income and African American mothers also value discussion of internal state and emotional language, and those
mothers who use more emotional language tend to have preschool-aged children with greater emotion situation knowledge (Blake, 1993; Garner, Jones, Gaddy, & Rennie, 1997).

Emotion situation knowledge may also be culture-specific in that certain situations may elicit different emotions depending on cultural expectations. Cole, Tamang, and Shrestha (2006) found that 3- to 5-year-old Nepali children were differentially socialized to display anger or shame in response to thwarted efforts, depending on whether they belonged to the high-caste Hindu Brahman group or the Buddhist Tamang group.

**EK and Behavioral Outcomes**

Not only do the components of EK emerge sequentially and at different points of childhood development, but each of the components of EK make different contributions to positive child outcomes. Researchers have found emotion situation knowledge to be a strong predictor of positive outcomes in young children, such as social functioning and school adjustment. (Denham, McKinley, Couchoud, & Holt, 1990; Garner, Jones, & Miner, 1994; Schultz et al., 2001). EK is not only associated with positive outcomes in children, but its acquisition in early childhood years is critical in social and emotional development (Astington, 1993; Cutting & Dunn, 1999; Denham & McKinley, 1993). It is important to have adequate measures of children’s EK during this critical developmental period because of the implications of its early development on later outcomes.

**Limitations of Current EK Measures**

Measures of EK have often examined the construct solely with the use of receptive EK tasks in which children match a spoken emotion descriptor or label to a picture of an emotion expression (Barth & Bastiani, 1997; Kusche, 1984; Mostow et al., 2002). In recent years, EK measures have been expanded to include one or more of the following: emotion matching, expressive EK, or emotion situation knowledge (Denham, 1986; Garner et al., 1994; Matsumoto, Yoo, Hirayama, & Petrova, 2005; Wang, 2003). EMTs require children to match photographs or pictures of like emotion expressions. Expressive EK tasks, however, require children to generate an emotion word that corresponds with either a cartoon drawing or a photograph of an emotion expression. Measures of emotion situation knowledge often utilize a series of vignettes in which provocative events occur (Denham, 1986; Wang, 2003). These vignettes are often accompanied by either puppets or cartoon pictures. Children must then match an emotion label or face to the depicted emotion-eliciting event.

Garner and colleagues have used a series of vignettes to measure emotion situation knowledge in young children. Children listen to 8 to 10 audiotaped vignettes depicting the emotions of happiness, sadness, anger, fear, and surprise. Children then select the best drawing of a facial expression to match the situation depicted in the vignette. This procedure can serve as a measure of emotion situation knowledge or can be combined with an expression labeling task in which children provide emotion labels for six drawings of facial expressions (Garner et al., 1994, 1997). A measure for older children is the diagnostic analysis of nonverbal accuracy scale (DANVA; Nowicki & Duke, 1994), which measures ability to send and receive non-verbal cues and includes subtests that measure receptive understanding of and expressive ability in creating facial expressions, gestures, postures, and tone of voice.
Past EK measures correlated highly with young children’s language ability. Cutting and Dunn (1999) found preschool-aged children’s scores on Denham’s (1986) AKT were highly positively correlated with their performance on two language measures, the British picture vocabulary scale (BPVS; Dunn, Dunn, Whetton, & Pintilie, 1997) and a narrative expressive language ability task called the ‘bus story’ (Renfrew, 1991). Due to the nature of EK measures, children with bigger vocabularies are able to comprehend the instructions with more ease, and thus, are able to perform better on the EK tasks. There remains a question as to whether EK measures adequately assess understanding of emotions rather than simply measuring an aspect of verbal ability. Thus, there is still need for an EK measure that can test the construct even after controlling for verbal ability.

**Affective Knowledge Test.** Perhaps one of the most widely used EK measures to date, Denham (1986)’s affective knowledge test (AKT or more widely known as the puppet interview) uses puppets with detachable faces to assess children’s understanding of emotion expressions and the situations that cause them. The AKT consists of an expressive EK task, receptive EK task, and two emotion situation knowledge tasks. In the stereotypical emotion situation knowledge task, children match emotion expression faces to a vignette depicting one of eight unequivocal emotion-provoking situations (i.e., getting an ice cream cone). In the non-stereotypical emotion situation knowledge task, children match emotion expression faces to 12 equivocal emotion-provoking situations.

AKT has strong correlates, including social competence and emotion regulation. Preschool-aged children’s EK as assessed by the AKT have related to higher social competence (Denham et al., 2003). The AKT has been shown to be affected by a socioemotional curriculum (promoting alternative thinking strategies; PATHS) for low-income preschool-aged children (Domitrovich, Cortes, & Greenberg, 2007).

**Kusche Emotional Inventory (KEI).** The KEI (Kusche, 1984) is a measure of receptive EK. Children are shown four cartoon pictures depicting emotion expressions and are asked to match an emotion label to one of the four pictures. Emotion labels include four basic emotions of happy, sad, mad, and scared, as well as the more complex emotions of confused, love, surprised, proud, disappointed, embarrassed, and tired. The KEI has been used to measure EK in preschool-aged children in several empirical studies. Higher EK scores on the KEI have been linked to better social functioning in children, specifically less negative peer nominations (Miller et al., 2005). Similar to the AKT, the KEI has also been shown to be affected by a socioemotional curriculum for low-income preschool-aged children (Domitrovich et al., 2007).

**The EMT: A New Measure of EK**

The EMT (Izard et al., 2003) was designed for preschool-aged children, including those who are economically disadvantaged. The EMT features brightly colored photographs of ethnically diverse children making facial expressions of happiness, sadness, anger, fear/surprise, and ‘neutral’ (no visible facial muscle movement that signals emotion). We chose to limit the EMT to these four basic emotions as understanding of these emotions emerges first in development, often during the preschool years (Denham, 2003). We have found in administering EMT that the use of these photographs makes the EMT highly attractive to ethnic minority children. Target
photographs (i.e., posed photographs of children representing the correct emotion for a given item) were standardized in a sample of 84 college students to meet a criterion of 80% agreement of emotion expression.

**Aspects of EK Measured by EMT.** The EMT consists of four parts that measure the components of receptive EK, emotion situation knowledge, and expressive EK. In the EMT, children match emotion expressions with expressions of the same category, with situations or causes, and with spoken emotion labels. Children also produce emotion labels for displayed photographs of emotion expressions. Thus, the EMT can measure different components of EK and serves as an efficient means of studying children’s ability to understand emotions in a variety of ways. Because young children develop receptive EK before acquiring expressive EK and because it is possible for young children to label emotion expressions but not understand which situations elicit those emotions (Denham, 2003; Odom & Lemond, 1972), it seemed useful to devise a way to efficiently examine the various components of EK. This is even more necessary because the components of EK have related differently with outcome variables (Garner et al., 1994). Furthermore, children 3 years old and younger or children from low-income backgrounds may not have acquired the emotion vocabulary for an expression labeling task or even an expression-label matching task, even though they may have mastered the ability to match emotion expressions of the same category. Adequate EK measures should include an expression-matching subtest as this task examines developmental progress in understanding emotions. Few, if any, existing EK measures contain this component.

**Preliminary Criterion Validity of the EMT.** The criterion validity of the EMT has been demonstrated in a study of an emotion-based preventive intervention program, the emotions course (EC). EC was designed to improve preschool-aged children’s EK, emotion regulation, and emotion utilization. Compared with a control group (Head Start as usual) and to a group receiving another socioemotional component (I can problem solve; ICPS; Shure, 2001), children who received EC improved their understanding of emotions as measured by an expression labeling task, an expression-label matching task (Izard, Trentacosta, King, & Mostow, 2004) and by the EMT (Izard et al., 2008). Thus, the EMT proved successful in measuring an intervention-induced increase in EK (Izard et al., 2008).

Children who participated in EC and improved their performance on the EMT, were also less likely to have teacher-rated levels of aggression, anxiety, and depression above clinical cutoffs. Also, children with higher scores on the EMT were also rated as having better emotion regulation by their teachers (Izard et al., 2008). Quite importantly, Izard et al. (2008) were able to demonstrate the EMT’s utility in a low-income, ethnically diverse population of preschoolers ranging in age from 3 years to 5 years.

**The Current Study**

In the current study, we examined the construct validity of the EMT by evaluating its criterion validity in predicting effortful control, emotion regulation, and behavioral outcomes. We evaluated the convergent validity of the EMT by comparing children’s performance on EMT to their performance on two widely used and previously validated measures of EK, the KEI and AKT (Denham, 1986; Kusche, 1984). We chose to compare the EMT to the KEI and AKT because they have been widely used with
preschool age children, including those in low-income families. Also, performance on both has been shown to be affected by a socioemotional curriculum (Domitrovich et al., 2007). In regard to criterion validity, we compared the EMT with KEI and AKT as predictors of emotion regulation, total childhood behavior problems, positive and negative interactions with peers and adults, and the temperament/personality characteristic of effortful control.

Method

Participants

The participants were 59 3- to 6-year-old children (range of 37 months to 74 months; SD 11.72 months) recruited from two university-setting daycare centers in the Mid-Atlantic region. We invited 54 children to participate from all three preschool classrooms in one of the daycare centers. The parents of 40 of these children gave consent for participation in our study. Thirty-eight of these children assented for participation. We invited 24 children from two of the preschool classrooms in the second daycare center to participate. We chose to invite only children from these two preschool classrooms as the other preschool classrooms in the center had received a socioemotional component as part of their standard curriculum. The parents of 21 of these children gave consent for participation in our study and all of these children assented for participation.

Of the 59 participants, 12 were unable to complete the AKT due to a missing parent report that is required for completion of this measure. Consequently, the N for this EK measure was smaller than that for EMT and KEI. Children’s mean age was 56.04 months. Approximately 69% were White, 16% were African American, 12% were Asian American, and 2% were Biracial/Other. Approximately 52% were males. We did not collect demographic information and thus have no information on their income level or SES.

Procedure

We administered three EK measures to preschool and kindergarten age children as well as a test of verbal ability, the Peabody picture vocabulary test. Individual testing sessions took place during times when children were most likely to be involved in free play or structured activities with their teachers. Research assistants administered the four measures to children in four testing sessions. Each test was administered during a separate testing session and on a different day to avoid fatigue. The testing order was PPVT, EMT, KEI, and AKT. Each testing session typically lasted for about 15 minutes and occurred in the classroom or within sight of it. Teachers and parents also filled out forms about their children’s regulation and related behaviors.

Measures

Verbal Ability. The Peabody picture vocabulary test, third edition (PPVT-III: Dunn & Dunn, 1997) was used to control for verbal ability across measures. In the PPVT-III, children matched a word to one of four pictures. Words were presented orally in increasing order of difficulty. The PPVT-III correlates highly with other measures of verbal ability and general intelligence (Bell, Lassiter, Matthews, & Hutchinson, 2001).
The PPVT-III has good internal consistency, $\alpha = .94$ (Dunn & Dunn, 1997). We used children’s standard scores on the PPVT-III as our measure of verbal ability.

**KEI.** We used a revised version of the recognition of emotion concepts subtest from the KEI (Kusche, 1984) as a means to evaluate the convergent validity of the EMT. This subtest, which consists of 30 items, is a receptive emotion vocabulary test which consists of matching emotion labels to cartoon drawings (Speltz, DeKlyen, Calderon, Greenberg, & Fisher, 1999). Fifteen emotions are presented two times each. These emotions consist of four basic emotions (happy, sad, mad, scared) as well as other, more complex emotions (e.g., disappointed, confused, embarrassed). Although the other two EK measures used in our study concern only four basic emotions, we chose to include these additional, more complex emotion labels in our administration of the KEI in order to ensure that that the KEI was administered as it was designed to be. Each set of four cartoon drawings per question includes one correct response, one incorrect response of the same valence (positive or negative), and two incorrect responses of the opposite valence. Similar to the Denham’s (1986) AKT, children received two points for a correct answer and one point for an incorrect answer of the same valence. The alpha for this measure was .86.

**Affective Knowledge Test.** We used Denham’s (1986) AKT, as a means to evaluate the convergent validity of the EMT. Children’s parents were asked to fill out a form indicating how their child would feel in a variety of situations from a choice of two emotions (e.g., either happy or sad when separating from parent to go to school). During individual testing sessions, children first completed an expressive EK task in which they labeled the four detachable emotion expression faces. Next, children completed a receptive EK task in which they matched four spoken emotion labels (happy, sad, mad, scared) with the four detachable emotion expression faces. Using faceless puppets, children then completed two emotion situation knowledge tasks. For the stereotypical emotion situation knowledge task, children watched the faceless puppets act out eight vignettes that display standard emotions (i.e., getting an ice cream cone or having a bad dream). Children then watched the faceless puppets act out 12 vignettes of equivocal emotion situations for the non-stereotypical emotion situation knowledge task. Based on the parent form, children watched puppets act out vignettes in which puppets display an emotion opposite of how the child would feel in the situation (i.e., meeting a very big, but friendly dog or separating from the parent when going to school). For both sets of vignettes, children were asked to attach the appropriate emotion face (happy, sad, mad, or scared) to the puppet. Children received two points for choosing the correct answer and one point for an incorrect answer of the same valence. The alpha for the total score on this measure was .87. The alphas for the expressive knowledge task, receptive knowledge task, stereotypical EK task, and non-stereotypical EK tasks were .21, .31, .64 and .81, respectively.

**EMT.** In the EMT (Izard et al., 2003), children respond to photographs of ethnically diverse elementary school children’s emotion expressions of happiness, sadness, anger, and surprise/fear. Ethnicities represented in the EMT stimuli include African American, Hispanic, White, and biracial. The EMT consists of four parts (matching expressions, expression-situation matching, expression labeling, and expression-label matching). In part 1 of the EMT (matching expressions), children match different expressions in the same emotion category. Children are asked to ‘show me which
one of these children (in a set of four photographs) feel the same way as this one (target photograph). In part 2 (expression-situation matching), children match expressions with situations or causes. For example, children are asked to ‘show me the one who got a pretty puppy for a birthday present’. In part 3 (expression labeling), children produce labels for the emotion expression of a single displayed picture. Children are asked to ‘look at her/his face. How does she/he feel?’ In part 4 (expression-label matching), children match emotion expressions and emotion labels. For example, children are asked to ‘show me the one who feels happy.’ In parts 1, 2, and 4, children are shown a quartet of pictures consisting of one target picture and three distracters. The quartet of pictures (including the target and distracters) changed for each item on parts 1, 2, and 4, for a total of 36 different quadrants presented in the EMT. Each of the four parts consists of 12 items. The alphas for each of the four parts were .65, .54, .76, and .80, respectively. The alpha for EMT total score (all four parts combined) is .88.

We examined the internal consistency of split half and alternate forms of the EMT. In the data, from the 59 3- to 6-year-old children, split half internal reliability coefficients for EMT-A and EMT-B were strong (α = .81 and .80, respectively). We created two 24-item forms of the EMT by dividing each part of the test in half. EMT A consisted of items 1, 2, 3, 4, 6, and 9 of part 1, the first half of part 2, items 1, 2, 3, 4, 5, and 8 of part 3, and the first half of part 4. EMT-B included items 5, 7, 8, 10, 11, and 12, the second half of part 2, items 6, 7, 9, 10, 11, and 12 of part 3, and the second half of part 4. We chose to divide the EMT this way in order to ensure that both forms of the test had a balance of four basic emotions. Reliability for the two 24-item forms was strong (α = .87).

Emotion Regulation. Children’s emotion regulation was rated by teachers using the emotion regulation checklist (ERC; Shields & Cicchetti, 1997). This measure contains 24 items assessing children’s affective intensity, emotional lability, and adaptive regulation. Each item is rated on a Likert scale ranging from 1 (rarely/never) to 4 (almost always), and there are two subscales, emotion regulation and lability/negativity. We used the emotion regulation subscale as our measure of emotion regulation in this study. The regulation subscale contains items such as ‘shows empathy; seems concerned when other people are upset or distressed’. The alpha of this subscale in our study was .64. In previous studies, the ERC has been used with preschool through elementary school children and consistently relates to children’s school and social adjustment outcomes (Shields, Dickstein, et al., 2001; Shields, Ryan, & Cicchetti, 2001).

Childhood Behavior Problems. Teachers filled out the caregiver-teacher rating form (C-TRF; Achenbach & Rescorla, 2000). The C-TRF is a standardized 100-item problem checklist. The C-TRF provides a measure of several childhood behaviors including externalizing and internalizing behaviors such as aggression, anxiety, and depression. We used the total behavior problems scale as our measure of childhood behavior problems. The CTRF has high subscale/syndrome alphas of .52–.96, with a total problems alpha of .97. Test-retest reliabilities are equally high. The C-TRF has content and criterion-related validity (Achenbach & Rescorla, 2000).

Independent Observers’ Ratings of Positive and Negative Interactions. Children were observed for 8-minute intervals during times when they were most likely to be involved
in free play or structured activities with their teachers. Research assistants coded the child’s behavior and emotion expression every 15 seconds within the 8-minute time period. Each child was coded for three 8-minute intervals. Children were coded for the number of adult and peer positive and negative interactions. Undergraduate coders were trained until they reached inter-rater reliability of .80.

Effortful Control. Parents filled out the very short form of the children’s behavior questionnaire (Rothbart, Ahadi, Hershey, & Fisher, 2001). This questionnaire asked parents to indicate how their child would react in a variety of situations. A sample item from the effortful control scale was ‘good at following instructions’. The alpha for this subscale was .83.

Results

Descriptive Data

Table 1 shows the means, standard deviations, and ranges for age, PPVT, EMT total and parts 1–4, KEI, and AKT and Table 2 presents the intercorrelations among the EK variables and the criterion variables.

Convergent Validity

Intercorrelations among EK Measures. Convergent validity coefficients were moderate to high among EMT total and EMT part 1–4 scores and KEI total score (rs = .70, .76, .56, .32, .63; ps < .10). Intercorrelations among EMT total and part 1–4 scores and AKT total score were also moderate to high (rs = .75, .65, .57, .49, .64; ps < .01).

Intercorrelations among EK Measures by EK Components. The receptive EK part of the EMT (EMT part 4) correlated strongly with overall KEI score (a receptive EK task; r = .71, p < .01). However, the receptive EK part of the EMT (EMT part 4) did not correlate significantly with the receptive EK part of the AKT. The expressive EK part of the EMT (EMT part 3) correlated moderately with the expressive EK part of the AKT (r = .36, p < .02). The emotion situation knowledge part of the EMT (EMT part 2) correlated highly with the stereotypical emotion situation knowledge part of the AKT (r = .56, p < .01) and with the non-stereotypical emotion situation knowledge part of the AKT (r = .48, p < .01).

Intercorrelations between EK Measures and Verbal Ability. EMT total correlated significantly with verbal ability (r = .56; p < .01). KEI total also correlated significantly with verbal ability (r = .52; p < .01). AKT correlated significantly with verbal ability (r = .47, p < .01). EMT parts 1–4 were also significantly correlated with verbal ability (rs = .45, .51, .38, .53, ps < .01). Fisher’s z-tests based on r to z transformation of correlation coefficients were conducted to determine whether the correlations of each EK measure with verbal ability differed significantly. The z statistic for EMT total and verbal ability versus KEI total and verbal ability was not significant (z = .28, p = .77). The z statistic for EMT total and verbal ability versus AKT and verbal ability was also not significant (z = .60; p = .55). AKT receptive EK task was significantly less correlated with verbal ability than EMT part 4 (z = -3.22, p < .001). Fisher’s r to z transformation and subsequent z-tests revealed no significant differences between
the correlations between the other AKT parts and age and the other EMT parts and age.

Intercorrelations between EK Measures and Age. EMT total correlated significantly with age (r = .59; p < .01). KEI total also correlated significantly with age (r = .63; p < .01). EMT parts 1–4 were also significantly correlated with age (rs = .67, .54, .46, .33; ps < .01). Again, Fisher’s z transformation and t-tests were conducted to determine whether these correlations significantly differed. EMT part 4 was significantly less correlated with age than EMT part 1 and KEI total (zs = 2.39, 2.01; ps < .05). AKT expressive EK was significantly less correlated with age than was EMT part 3 and KEI total (zs = −2.40; −3.57; ps < .01). Fisher’s r to z transformation and subsequent z-tests revealed no significant differences between the correlations of EMT total and age and AKT total and age.
Table 2. Intercorrelations among Age, Verbal Ability, EK Measures and Criterion Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMT total</td>
<td>.59**</td>
<td>.56**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEI part 1</td>
<td>.67**</td>
<td>.45**</td>
<td>.82**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEI part 2</td>
<td>.54**</td>
<td>.51**</td>
<td>.86**</td>
<td>.72**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEI part 3</td>
<td>.46**</td>
<td>.38**</td>
<td>.82**</td>
<td>.49**</td>
<td>.59**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEI part 4</td>
<td>.33**</td>
<td>.53**</td>
<td>.84**</td>
<td>.59**</td>
<td>.60**</td>
<td>.59**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEI total</td>
<td>.63**</td>
<td>.52**</td>
<td>.82**</td>
<td>.79**</td>
<td>.70**</td>
<td>.58**</td>
<td>.71**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKT</td>
<td>.28'</td>
<td>.47**</td>
<td>.56**</td>
<td>.51**</td>
<td>.49**</td>
<td>.38**</td>
<td>.58**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKT rec</td>
<td>.09</td>
<td>-.06</td>
<td>.13</td>
<td>.19</td>
<td>-.01</td>
<td>.12</td>
<td>.12</td>
<td>.46**</td>
<td>.48**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKT exp</td>
<td>.01</td>
<td>.35*</td>
<td>.32*</td>
<td>.23</td>
<td>.29*</td>
<td>.36*</td>
<td>.16</td>
<td>.22</td>
<td>.61**</td>
<td>.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKT ster</td>
<td>.32*</td>
<td>.53**</td>
<td>.54**</td>
<td>.50**</td>
<td>.56**</td>
<td>.42**</td>
<td>.34*</td>
<td>.52**</td>
<td>.69**</td>
<td>.12</td>
<td>.25'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKT non</td>
<td>.28'</td>
<td>.42**</td>
<td>.55**</td>
<td>.54**</td>
<td>.47**</td>
<td>.44**</td>
<td>.38**</td>
<td>.53**</td>
<td>.95**</td>
<td>.37**</td>
<td>.53**</td>
<td>.51**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERC reg</td>
<td>.44**</td>
<td>.20</td>
<td>.43**</td>
<td>.35*</td>
<td>.37**</td>
<td>.31*</td>
<td>.45**</td>
<td>.41**</td>
<td>.10</td>
<td>-.03</td>
<td>-.01</td>
<td>.15</td>
<td>.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTRF</td>
<td>.32*</td>
<td>-.32*</td>
<td>-.32*</td>
<td>-.30*</td>
<td>-.37**</td>
<td>-.22</td>
<td>-.21</td>
<td>-.29</td>
<td>-.35*</td>
<td>.14</td>
<td>.08</td>
<td>-.20</td>
<td>-.25</td>
<td>-.28*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pos inter</td>
<td>.27'</td>
<td>.13</td>
<td>.26'</td>
<td>.17</td>
<td>.23</td>
<td>.23</td>
<td>.20</td>
<td>.38**</td>
<td>.19</td>
<td>.08</td>
<td>-.20</td>
<td>-.15</td>
<td>-.61**</td>
<td>.22</td>
<td>-.27'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neg inter</td>
<td>.09</td>
<td>-.43**</td>
<td>-.30*</td>
<td>-.14</td>
<td>-.29*</td>
<td>-.21</td>
<td>-.31*</td>
<td>-.19</td>
<td>-.47**</td>
<td>-.02</td>
<td>-.10</td>
<td>-.15</td>
<td>-.42**</td>
<td>.02</td>
<td>.55**</td>
<td>-.11</td>
<td></td>
</tr>
<tr>
<td>Eff ctrl</td>
<td>.00</td>
<td>.15</td>
<td>.29'</td>
<td>.10</td>
<td>.29*</td>
<td>.33*</td>
<td>.18</td>
<td>.18</td>
<td>.16</td>
<td>-.22</td>
<td>.09</td>
<td>.16</td>
<td>.26</td>
<td>.08</td>
<td>-.42**</td>
<td>.21</td>
<td>-.29*</td>
</tr>
</tbody>
</table>

Note: ** p < .01; * p < .05; ' p < .10.
Criterion Validity

Correlations with Other Measures. Criterion validity coefficients were strong among EMT total and EMT part 1–4 scores and measures of teacher-rated emotion regulation, teacher-rated childhood behavior problems, observed adult and peer positive and negative interactions, and parent-rated effortful control. EMT total score correlated positively with the regulation subscale of the ERC \( (r = .45, p < .001) \). Parts 1–4 of the EMT also correlated positively with the regulation subscale of the ERC \( (rs = .35, .37, .31, .45; ps < .03) \). EMT total score significantly correlated with teacher-rated total childhood behavior problems (CTRF Total score, \( r = -.32, p < .02 \)). Parts 1 and 2 of EMT also correlated with total childhood behavior problems \( (r = -.30, p < .03; r = -.37, p < .01) \). EMT total score significantly correlated with observed adult and peer negative interactions \( (r = -.30; p < .05) \). Only parts 2 and 3 of EMT significantly correlated with negative interactions \( (r = -.29; p < .05; r = -.31, p < .04) \). Only parts 2 and 3 of EMT significantly correlated with parent reported effortful control \( (r = .29; p < .05; r = .33, p < .02) \). Correlations between the criterion variables and KEI total and AKT are included in Table 2. Fisher’s z transformations were run on each set of correlations to determine if the correlations between the criterion variables and the EK measures differed significantly. None of the calculated \( z \) statistics were significant.

Regression Analyses

We conducted block entry regression analyses to examine the relations of the EMT and our comparison EK measures to our criterion variables of teacher-rated emotion regulation, teacher-rated total childhood behavior problems, observed adult/peer positive and negative interactions, and parent reported effortful control. We first evaluated EMT as a predictor of our criterion variables. For each of these models, age, sex, PPVT score, and EMT total were entered as predictor variables in that order. We next evaluated the KEI as a predictor of our criterion variables. For each of these models, age, sex, PPVT score, and KEI were entered as predictor variables in that order. We next evaluated AKT as a predictor of each of our criterion variables. For each of these models, age, sex, PPVT score, and AKT were entered as predictor variables in that order. Tables 3–5 provide values derived from the full model of each of the significant regression analyses.

Table 3. The Regression of Effortful Control on Age, Sex, PPVT, and EMT Total

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.02</td>
<td>.01</td>
<td>-.26</td>
</tr>
<tr>
<td>Sex</td>
<td>-.26</td>
<td>.21</td>
<td>-.19</td>
</tr>
<tr>
<td>PPVT</td>
<td>-.01</td>
<td>.01</td>
<td>-.33</td>
</tr>
<tr>
<td>EMT total</td>
<td>-.05</td>
<td>.02</td>
<td>.62*</td>
</tr>
</tbody>
</table>

*Note: The values in the table come from the final model of the regression analysis.

\* \( p < .05 \).

\( \Delta R^2 \) for each variable: Age = .00; Sex = .02; PPVT = .00; EMT Total = .15*.
For our analysis in which CBQ effortful control was the criterion variable, total EMT score significantly predicted effortful control ($\beta = .62$, $p < .02$; Table 3). EMT score did not significantly predict the regulation scale of the ERC, observed adult/peer positive interactions, observed adult/peer negative interactions, total childhood behavior problems, or parent reported effortful control.

**KEI and AKT.** KEI score did not significantly predict teacher-rated emotion regulation, observed adult/peer positive or negative interactions, childhood behavior problems, or parent reported effortful control. AKT score did not significantly predict teacher-rated emotion regulation, observed adult/peer positive or negative interactions, childhood behavior problems, or parent reported effortful control.

**Exploratory Analyses**

We conducted exploratory analyses to determine whether EMT parts 1–4 contributed differentially to the prediction of our child outcomes. For each of our analyses, we

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>$SE$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.01</td>
<td>.01</td>
<td>-.15</td>
</tr>
<tr>
<td>Sex</td>
<td>-.28</td>
<td>.22</td>
<td>-.20</td>
</tr>
<tr>
<td>PPVT</td>
<td>-.01</td>
<td>.01</td>
<td>-.19</td>
</tr>
<tr>
<td>EMT part 2</td>
<td>.13</td>
<td>.06</td>
<td>.48*</td>
</tr>
</tbody>
</table>

*Note: The values in the table come from the final model of the regression analysis.  
$* p < .05$.  
$\Delta R^2$ for each variable: Age = .00; Sex = .02; PPVT = .01; EMT Part 2 = .12*.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>$SE$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.08</td>
<td>.03</td>
<td>.30</td>
</tr>
<tr>
<td>Sex</td>
<td>-.43</td>
<td>.74</td>
<td>-.07</td>
</tr>
<tr>
<td>PPVT</td>
<td>-.02</td>
<td>.03</td>
<td>-.09</td>
</tr>
<tr>
<td>EMT Part 4</td>
<td>-.36</td>
<td>.14</td>
<td>.41*</td>
</tr>
</tbody>
</table>

*Note: The values in the table come from the final model of the regression analysis.  
$* p < .05$.  
$\Delta R^2$ for each variable: Age = .20**; Sex = .00; PPVT = .02; EMT Part 4 = .10*.
entered age, sex, PPVT and EMT parts 1, 2, 3, and 4 as predictor variables, in that order. Part 2 of the EMT, expression-situation matching, significantly predicted children’s effortful control ($\beta = .48, p < .03$; Table 4). Part 4 of the EMT, label-emotion expression matching, significantly predicted emotion regulation score ($\beta = .41, p < .02$; Table 5). We conducted exploratory analyses to determine whether the AKT parts that correspond to EMT parts 2 and 4 also significantly predicted our child outcomes. Neither AKT receptive EK task nor AKT stereotypical emotion situation knowledge task significantly predicted our child outcome variables.

**Discussion**

The EMT correlated strongly with two widely used measures of EK, the KEI and AKT. The four individual parts of the EMT, matching expressions, expression-situation matching, expression labeling, and expression-label matching, also correlated strongly with the KEI and the parts of the AKT. Part 2 of the EMT, expression-situation matching, correlated strongly with both emotion situation knowledge parts of the AKT. Part 3 of the EMT, expression labeling, correlated strongly with the expressive EK part of the AKT. Part 4 of the EMT, expression-label matching, did not correlate significantly with the receptive EK part of the AKT but did correlate significantly with the KEI, a receptive emotion vocabulary test.

Because evidence shows that the KEI and the AKT accurately assess children’s EK and adequately predict correlates of EK such as emotion regulation and social competence (Denham et al., 2003; Miller et al., 2005; Suveg, Kendall, Comer, & Robin, 2006), the strong correlations between the EMT and these two measures provide solid support for the EMT’s convergent validity, key to overall construct validity.

One limitation of EK measures as predictors of behavioral outcomes is their high correlation with age and verbal ability. KEI, AKT and EMT total scores correlated similarly with age and verbal ability. AKT receptive EK task was less strongly correlated with verbal ability than was part 4 of the EMT. AKT expressive EK task was less strongly correlated with age than was part 3 of the EMT. Thus, we were unable to demonstrate that the EMT is less related to verbal ability and age than the other two EK measures.

However, part 4 of the EMT was less correlated with age than KEI total and EMT part 1 were. This finding lends some support to the use of part 4 of the EMT for younger age children as a measure of EK. This finding also supports the developmental literature as EMT part 4 tests children’s receptive EK, which tends to emerge first in development (Denham, 2003). Still, it seems prudent always to control for age and verbal ability for the best estimate of EK as a determinant of aspects of behavioral outcomes.

Fisher’s $z$-test based on $r$ to $z$ transformation of scores also demonstrated that the correlations of EMT with behavioral outcomes are similar to those of other measures of EK with the same behavioral outcomes. Specifically, EMT correlated moderately (and positively) with emotion regulation and moderately but negatively with childhood behavior problems and observed adult and peer negative interactions. EMT’s positive correlation with observed adult and peer positive interactions approached significance. Overall, these results lend support to the EMT’s criterion validity.

The EMT was a more robust predictor of emotion regulation and effortful control than were the KEI and AKT. Particularly, part 4 of the EMT, expression-label matching, predicted higher emotion regulation. Thus, children who can match a spoken
emotion label with one of four emotion expressions are better able to regulate their emotions. This finding is consistent with differential emotions theory in showing that understanding of emotions contributes to greater ability to modulate emotion states (Izard et al., 2008; Lieberman et al., 2007).

Furthermore, overall EK (EMT total score) was related to children’s ability to control their actions (i.e., effortful control). In examining the contribution of the individual parts of EMT, part 2, expression-situation matching, emerged as the EK component driving this prediction. This finding also fits differential emotions theory in that children who can understand which events are related to which emotions show greater ability to cognitively control their actions (i.e., effortful control; Izard, 2002). It is possible that this relation is bidirectional and that children who are high on effortful control may also have a greater ability to sustain attention and thereby attain greater EK and higher scores on the EMT. Nevertheless, higher scores on effortful control have been associated with higher academic competence in children. Thus, our finding that the EMT predicts effortful control further supports EK’s positive relation with academic competence in young children (Fabes et al., 2003).

In the current study, the EMT did not significantly predict three of our five criterion variables, childhood behavior problems, observed positive adult/peer interactions, or observed negative adult/peer interactions. The KEI and AKT failed to predict all five criterion variables. They did not predict effortful control or emotion regulation, two criterion variables that the EMT did predict in this study. As these two EK measures have been found to relate to similar behavioral outcomes in previous studies (Denham et al., 2003; Miller et al., 2005), the lack of relation between the EK measures and the criterion variables in this study may be due to low power. Evidence of the EMT’s criterion validity should be considered preliminary and more research with a larger N is needed to evaluate its criterion validity.

**EMT Total vs. EMT Parts**

The EMT is designed to be administered as a relatively comprehensive (four-part) measure of EK in young children. However, our findings suggest that part 4 of the EMT may be useful as a brief measure of EK for younger age children, particularly for predicting their emotion regulation ability. There is also some evidence that part 2 of the EMT may be used to predict effortful control. However, more research is needed to confirm that these parts differentially predict these two important behavioral outcomes.

**Strengths of the EMT**

The EMT serves as an EK measure that assesses preschool children’s emotion matching (part 1 of the EMT, matching different expressions in the same emotion category), emotion situation knowledge (part 2 of the EMT, matching expressions and situations/causes of emotion), expressive EK (part 3 of the EMT, producing labels for emotion expressions), and receptive EK (part 4 of the EMT, matching emotion expressions and emotion labels that were spoken by the examiner). Thus similar to the AKT, the EMT serves as a comprehensive measure of EK. It also enables the researcher to determine the relative strength of each of the four components of EK in driving the relations between overall EK and its correlates. The EMT was a more robust predictor of emotion regulation and effortful control than were the AKT and KEI.
Although parts of the AKT were less correlated with age and verbal ability than were the parts of the EMT, part 4 of the EMT was less correlated with age than the KEI and the other parts of the EMT, indicating that part 4 of the EMT may be useful as a brief measure of EK for young children. Further research with larger samples is needed to examine how each of the four components of EK represented in the EMT relates to children’s behavioral outcomes.

Limitations of the EMT

A computerized version of the EMT or a version that uses manipulatives might shorten the procedure as a result of more focused attention in the preschool participants. However, the use of such a computerized version or manipulatives may decrease the ease and standardization of training and administration of the EMT.

The length of the EMT also may create a disadvantage to young children who are just beginning preschool and those who have poor attention or delayed cognitive development. As our split half reliability coefficients for EMT A and EMT B are strong, it may be appropriate to use a 24 item version of the EMT for younger preschoolers, especially when time available for measuring EK is restricted.

Limitations of the Study

Administering three EK measures could have resulted in practice effects for measures in the second and third position in the order of testing. Because we wanted to focus on the construct and criterion validity of the new measure of EK, we eliminated practice effects on the EMT by administering it first and the KEI and AKT second and third respectively. However, this order may have inflated the EK scores for the KEI and AKT. The children may have gained some emotion vocabulary and become more familiar with test content and test-taking procedures during the administration of EMT. To prevent fatigue, each measure was administered in a separate testing session and on a different day.

Although evidence for the EMT’s convergent validity was strong, the EMT did not predict three of the five criterion variables in the current study. Past research has shown performance on the EMT to predict children’s positive and negative behaviors, including internalizing behaviors and externalizing behaviors (Izard et al., 2004, 2008). More research is needed to demonstrate the criterion validity of the EMT.

Implications for Research

Similar to the AKT and KEI, the EMT has been shown to be affected by a socioemotional curriculum designed for low-income preschool-aged children. The EMT can be used to evaluate the effectiveness of emotion-based preventive interventions, such as EC. Children who received EC had higher scores on the EMT and other positive outcomes. This relation provided evidence of criterion validity of the EMT and the current study provided evidence of the construct validity of the EMT. An implication of the findings from our research on the role of EK in preventive intervention is that the EMT may also prove useful in clinical research concerned with emotion-related psychological disorders in young children.
References


Acknowledgments

We thank the staff and teachers and the parents and children of the University of Delaware’s Early Learning Center and Laboratory Preschool. Their co-operation made this research possible. Thanks to Fran Haskins, Jenny Anderson, Stephanie Krauthamer-Ewing, and Kristy Sheffler for help with many aspects of the research. This research was supported by University of Delaware’s Graduate Scholars Award to the first author and by grants # 1 R21 MH068443-01A1 and # 1 R01 MH80909-01A1 from NIMH, Carroll Izard, PI.