Measurement of emotion recognition ability and emotional understanding
• Emotion recognition ability (ERA)
• How can ERA be measured & problems of existing tests
• Ways to improve the measurement of ERA
• The Geneva Emotion Recognition Test (GERT)
  – Development
  – Factor structure
  – Difficulty & measurement precision (IRT)
  – Construct validity
  – Predictive validity: negotiation performance
• Summary
Emotion recognition ability (ERA) =

The ability to accurately recognize the type and intensity of others’ emotional states from their nonverbal expressions conveyed by the face, voice, body (Elfenbein et al., 2007)
ERA allows to...
- anticipate another person’s actions
- adapt one’s own actions accordingly
  - smooth interactions
  - attain individual goals

People with high ERA...
- are rated as more likable, socially supportive, honest, open, and less hostile by their peers
- have more close relationships and feel less isolated and frustrated in them
- Achieve higher academic and workplace performance

Deficits in ERA are associated with...
- Antisocial behavior (children and adults)
- Mental disorders (schizophrenia, depression, Parkinson’s disease, eating disorders...)

ERA tests are used in many fields of psychology and have gained more attention recently as a possible alternative to emotional intelligence tests

⇒ Accurate and valid measurement of individual differences in ERA is needed

Hampson et al., 2006; Hall et al., 2005 & 2009; Carton et al., 1999; Kang, 2012; Elfenbein et al., 2007; Marsh & Blair, 2008; Cherniss, 2010
ERA tests use a forced choice response format with one correct answer:

**Please choose the term which best describes the emotion expressed in the preceding photo.**

- ANGER
- FEAR
- DISGUST
- SADNESS
- JOY

ERA is calculated as the sum of correctly recognized stimuli.
<table>
<thead>
<tr>
<th>Test</th>
<th>N emotions (N positive)</th>
<th>Length (min)</th>
<th>Face</th>
<th>Voice</th>
<th>Gestures/body</th>
<th>Multimodal stimuli</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>PONS (1979) [MiniPONS, 2011]</td>
<td>20 (10) affective states</td>
<td>45 [15]</td>
<td>x (dynamic)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>DANVA (1994)</td>
<td>4 (1)</td>
<td>15</td>
<td>x (still)</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JACBART (2000)</td>
<td>5 (1)</td>
<td>15</td>
<td>x (still)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSCEIT (2003) (subtest)</td>
<td>5 (1)</td>
<td>10</td>
<td>x (still + abstract pictures)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MERT (2009)</td>
<td>10 (2)</td>
<td>45</td>
<td>x (still &amp; dynamic)</td>
<td>x</td>
<td>(x)</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>ERI (2010)</td>
<td>5 (1)</td>
<td>10</td>
<td>x (still)</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Problems of ERA tests

• **Restricted ecological validity**
  - Many tests use stimuli from single modalities
  - Many tests contain only few emotions: discrimination or recognition?

• **Unclear psychometric quality**
  - Difficulty is low (tests are easy) → limited discrimination between subjects
  - Low internal consistency (Cronbach’s alpha) → one reason for why correlations between tests are low
  - Unknown factor structure: different skills for different emotions/ modalities or one general ERA?

Main obstacle: One needs emotion expressions that are difficult enough to discriminate between people, but it is tricky to find difficult portrayals that still express the intended emotion well enough
Ways of improving ERA measurement

- **Ecological validity:**
  - multimodal and dynamic portrayals, large number of affective states, spontaneous expressions

- **Difficulty and internal consistency:**
  - Use additional criteria for selecting not too easy but still recognizable portrayals (e.g., ratings of believability, FACS coding)

- **Use of appropriate statistical techniques:**
  - Item Response Theory: allows to evaluate measurement precision of each item and to adapt test difficulty to the ERA level of the population of interest
  - possibility to test for differential item functioning (e.g., gender bias)
  - Specific internal consistency indices for binary data
  - Exploratory and confirmatory factor analysis for binary data

- **Validation to meet psychometric quality standards:**
  - Construct validity/ nomological network
  - Predictive validity: Predict theoretically expected outcome variables
GEMEP database (Bänziger et al., 2010)
1260 video clips
• 10 actors
• 18 emotions
• 4 intensities
• 2 verbal contents

Selection of first item pool (based on ratings by N=20)
108 video clips
• 14 emotions, normal intensity
• verbal content = pseudolinguistic sentences
• target emotion must be most frequently chosen category
• believability rating & recognition rate must be > 30th percentile of all portrayals for an emotion
• male and female portrayals

Final GERT (based on study with N=295)
83 video clips
• 6 per emotion (3 male/female)
• Combination of factor analysis and Rasch model (Item Response Theory)
Confirmatory factor analyses on the 14 GERT emotion scores were performed (N=295). Best fitting model: Two highly correlated factors for positive & negative ERA with residual correlations

ERA is essentially unidimensional. It encompasses distinguishable, but correlated skills for specific emotion families and positive & negative emotions.

Rasch model results, N=295

Distribution of person parameters (ability scores)

Distribution of item parameters (difficulties)

Alpha: .76; Test score reliability (Dimitrov, 2003): .92

### Construct Validity

<table>
<thead>
<tr>
<th>Variable</th>
<th>Correlation with GERT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (1=female)</td>
<td>.13*</td>
</tr>
<tr>
<td>Age</td>
<td>-.49**</td>
</tr>
<tr>
<td>Multimodal Emotion Recognition Test (MERT)</td>
<td>.52**</td>
</tr>
<tr>
<td>Japanese &amp; Caucasian Affect Recognition Test (JACBART)</td>
<td>.49**</td>
</tr>
<tr>
<td>Mini Profile of Nonverbal Sensitivity (MiniPONS)</td>
<td>.43**</td>
</tr>
<tr>
<td>General Reasoning</td>
<td>.38**</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>.36**</td>
</tr>
<tr>
<td>Big Five - Neuroticism</td>
<td>.28**</td>
</tr>
<tr>
<td>Big Five - Openness</td>
<td>.30**</td>
</tr>
<tr>
<td>Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT)</td>
<td>.47**</td>
</tr>
<tr>
<td>Situational Test of Emotion Understanding (STEU)</td>
<td>.50**</td>
</tr>
<tr>
<td>Trait Emotional Intelligence Questionnaire (TEIQUE)</td>
<td>-.09</td>
</tr>
</tbody>
</table>

### Study Details

- **Study 1**
  - N=295
  - Male = 82

- **Study 2**
  - N=131
  - Male = 67

- **Study 3**
  - N=121
  - Male = 47
Are people scoring high on the GERT more successful in interpersonal situations?

→ Negotiation between a „recruiter“ and an „employee“ about a job contract (Pinkley et al., 1994)

<table>
<thead>
<tr>
<th>Work start</th>
<th>options</th>
<th>Employee</th>
<th>Recruiter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>June 1</td>
<td>2,400</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>June 15</td>
<td>1,800</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td>July 1</td>
<td>1,200</td>
<td>1,200</td>
</tr>
<tr>
<td></td>
<td>July 15</td>
<td>600</td>
<td>1,800</td>
</tr>
<tr>
<td></td>
<td>August 1</td>
<td>0</td>
<td>2,400</td>
</tr>
</tbody>
</table>

Goal: To win as many points as possible

**Hypotheses:**

• Higher GERT scores predict higher individual and joint gains.
• Negotiators with higher GERT scores are perceived as more cooperative.
2. Individual gains:

3. Cooperativeness/ likability rated by negotiation partner:

N=130 (65 same-gender dyads, 33 male)

1. In male dyads, recruiters’ GERT scores predicted higher joint gains for the dyad.

High GERT scores predicted objective and subjective performance mainly for the recruiter (high power) role. This was not found for cognitive intelligence. Results for Emotional Intelligence were inconsistent.
Limitation: Language differences

Some emotions are better recognized in French

Some emotions are better recognized in German

<table>
<thead>
<tr>
<th></th>
<th>Mean French (N=131)</th>
<th>Mean German (N=295)</th>
</tr>
</thead>
<tbody>
<tr>
<td>anger</td>
<td>0.74</td>
<td>0.55</td>
</tr>
<tr>
<td>disgust</td>
<td>0.73</td>
<td>0.50</td>
</tr>
<tr>
<td>fear</td>
<td>0.67</td>
<td>0.47</td>
</tr>
<tr>
<td>surprise</td>
<td>0.62</td>
<td>0.42</td>
</tr>
<tr>
<td>sadness</td>
<td>0.60</td>
<td>0.80</td>
</tr>
<tr>
<td>despair</td>
<td>0.47</td>
<td>0.72</td>
</tr>
<tr>
<td>total</td>
<td>0.70</td>
<td>0.67</td>
</tr>
</tbody>
</table>

- Difficult (impossible?) to find translations for emotion words that have the same meaning across languages
- Total accuracy does not differ between languages
- Psychometric properties hold in French and German
- We found evidence for construct validity in German and French

→ GERT seems to measure individual differences in ERA within each language, but test results might not be comparable across cultures
• higher ecological validity than previous tests → a broader measure of «emotion perception» dimension of EI (?)

• Essentially undimensional factor structure, but evidence for more specific subcomponents

• satisfactory measurement properties, but language/cultural differences

• good construct validity

• predictive validity in a face-to-face interaction (negotiation)

• Ongoing/future projects: training ERA, neural basis of ERA, predictive validity of ERA in nurses, managers...
Thank you for your attention!

links to other ERA tests: www.affective-sciences.org/webexperimentation
latest updates on the GERT: www.affective-sciences.org/gert
Questions? MarcelloMortillaro@unige.ch , Katja.Schlegel@unige.ch
Please select the emotion word that describes best the emotion that the actor tried to express in this video.
Construct validity

Geneva Emotion Recognition Test
1
Multimodal Emotion Recognition Test
2
Emotion Recognition Index
3
Brief Profile of Nonverbal Sensitivity
4
Japanese and Caucasian Brief Affect Recognition Test
5
Mayer-Salovey-Caruso Emotional Intelligence Test
6
Situational Test of Emotional Understanding
7
Situational Test of Emotion Management
8
General Reasoning (NV5-R)
9
Attention (NV5-R)
10
Vocabulary (NV5-R)
11
Extraversion (BFI)
12
Agreeableness (BFI)
13
Conscientiousness (BFI)
14
Openness (BFI)
15
Empathy (IRI)
16
Trait Emotional Intelligence Questionnaire
17
Rosenberg Self-Esteem Scale
18
Adaptive Functioning (ASR)
19
Personal Strength (ASR)
20
Neuroticism (BFI)
21
Trait Anxiety (STAIT)
22
Trait anger (STAXI)
23
Personal Distress (IRI)
24
Toronto Alexithymia Scale
25
Depressive symptoms (ASR)
26
Anxiety symptoms (ASR)
27
Withdrawal symptoms (ASR)
28
Somatic symptoms (ASR)
29
Thought problems (ASR)
30
Attention problems (ASR)
31
Aggressive behavior (ASR)
32
Norm-violating behavior (ASR)
33

Emotion recognition tests

Emotional intelligence tests

Cognitive intelligence tests

Positive personality traits & adjustment

Negative personality traits & syndromes