I know how you feel

Emotion recognition accuracy and training in psychotherapy education

Lillian Döllinger
I know how you feel
Emotion recognition accuracy and training in psychotherapy education
Lillian Döllinger

Academic dissertation for the Degree of Doctor of Philosophy in Psychology at Stockholm University to be publicly defended on Friday 27 October 2023 at 10.00 in hörsal 4, hus 2, Albano, Albanovägen 18.

Abstract
Researchers, practitioners and legislators agree that it is important to understand which kinds of psychotherapeutic treatments lead to lasting positive changes in patients’ well-being, how those treatments can be administered in efficient ways and how it can be determined which patients would benefit from which treatment. In recent years, there has also been growing interest in those who practice psychotherapy; specifically, in the socio-emotional and interpersonal characteristics and competencies that psychotherapists should possess to provide high quality treatments for a variety of patients, irrespective of psychotherapy approach. This thesis studies one such important psychotherapist competency, namely the ability to recognize non-verbal emotional expressions in others. Psychotherapists need to be able to help patients experience, understand and express their emotions, and, in this context, it is crucial that they themselves have good socio-emotional competencies, like emotion recognition accuracy. Still, there is surprisingly little research about psychotherapists’ emotion recognition accuracy and about how they could be supported in improving this ability in the course of their education.

Study I explores trainee psychotherapists’ emotion recognition accuracy in the beginning and in the end of theoretical and practical psychotherapy education, and compares it to a control group of undergraduate students. The results reveal that trainee psychotherapists in the beginning of their education show superior emotion recognition accuracy for multimodal (audio, video, audio-video) emotional expressions and micro expressions (<200ms) compared to the control group. This suggests that those who choose to become psychotherapists might already possess elevated emotion recognition accuracy or might have developed it early on during their studies. However, after one and a half years of education, their multimodal and micro expression emotion recognition accuracy does not improve significantly more than the control groups’ accuracy. This suggests that standard (psychodynamic and cognitive behavioral) psychotherapy education does not automatically lead to improved emotion recognition accuracy, even though the trainees learn how to conduct psychotherapy and also treat their first patients at the university clinic. Or, alternatively, that the socio-emotional competencies that develop during the education might not be captured by the standardized computerized emotion recognition accuracy tasks used in this study. Nonetheless, this finding might also suggest that more explicit training of emotion recognition accuracy is needed.

Study II then investigates two newly developed standardized computerized emotion recognition accuracy trainings, one for multimodal emotion recognition accuracy and one for micro expression recognition accuracy. The trainings are evaluated in a sample of undergraduate students using a mixed design. The trainings are compared to one another and to an active control training. Both trainings are found to significantly improve the participants’ emotion recognition accuracy in a one-week posttest. Study III extends those findings using a sample of trainee psychotherapists. Also in the target population, both trainings are found to be effective in the one-week posttest. In addition to that, the multimodal training shows effects for unimodal emotion recognition accuracy for audio-only and video-only stimuli in the one-year follow-up towards the end of psychotherapy education. This indicates that standardized computerized emotion recognition accuracy training can be used as a tool for improving trainee psychotherapists’ emotion recognition accuracy, even though additional interventions might be needed for securing long-term success for all facets of emotion recognition. Future research should explore the practical impact of trainee psychotherapists’ emotion recognition accuracy and the training of this ability.

The findings of this thesis are, on the one hand, surprising, in so far as psychotherapy education likely does not lead to improvements in trainee psychotherapists’ emotion recognition accuracy. On the other hand, they are encouraging, because they suggest that this ability can be trained with relatively simple and resource-efficient methods. Emotion recognition accuracy training could become part of standard or individualized psychotherapy training, alongside the training of other relevant verbal and non-verbal socio-emotional and interpersonal psychotherapist competencies.

Keywords: emotion recognition accuracy, psychotherapy education, trainee psychotherapists, psychotherapist competencies, psychotherapist characteristics, emotion recognition accuracy training, multimodal emotion recognition, micro expressions, psychodynamic psychotherapy, cognitive behavioral therapy.
I KNOW HOW YOU FEEL

Lillian Döllinger
I know how you feel
Emotion recognition accuracy and training in psychotherapy education
Lillian Döllinger
To my mother.
Researchers, practitioners and legislators agree that it is important to understand which kinds of psychotherapeutic treatments lead to lasting positive changes in patients’ well-being, how those treatments can be administered in efficient ways and how it can be determined which patients would benefit from which treatment. In recent years, there has also been growing interest in those who practice psychotherapy; specifically, in the socio-emotional and interpersonal characteristics and competencies that psychotherapists should possess to provide high quality treatments for a variety of patients, irrespective of psychotherapy approach. This thesis studies one such important psychotherapist competency, namely the ability to recognize non-verbal emotional expressions in others. Psychotherapists need to be able to help patients experience, understand and express their emotions, and, in this context, it is crucial that they themselves have good socio-emotional competencies, like emotion recognition accuracy. Still, there is surprisingly little research about psychotherapists’ emotion recognition accuracy and about how they could be supported in improving this ability in the course of their education.

**Study I** explores trainee psychotherapists’ emotion recognition accuracy in the beginning and in the end of theoretical and practical psychotherapy education, and compares it to a control group of undergraduate students. The results reveal that trainee psychotherapists in the beginning of their education show superior emotion recognition accuracy for multimodal (audio, video, audio-video) emotional expressions and micro expressions (<200ms) compared to the control group. This suggests that those who choose to become psychotherapists might already possess elevated emotion recognition accuracy or might have developed it early on during their studies. However, after one and a half years of education, their multimodal and micro expression emotion recognition accuracy does not improve significantly more than the control groups’ accuracy. This suggests that standard (psychodynamic and cognitive behavioral) psychotherapy education does not automatically lead to improved emotion recognition accuracy, even though the trainees learn how to conduct psychotherapy and also treat their first patients at the university clinic. Or, that the socio-emotional competencies that develop during the education might
not be captured by the standardized computerized emotion recognition accuracy tasks used in this study. Nonetheless, this finding might also suggest that more explicit training of emotion recognition accuracy is needed. **Study II** then investigates two newly developed standardized computerized emotion recognition accuracy trainings, one for multimodal emotion recognition accuracy and one for micro expression recognition accuracy. The trainings are evaluated in a sample of undergraduate students using a mixed design. The trainings are compared to one another and to an active control training. Both trainings are found to significantly improve the participants’ emotion recognition accuracy in a one-week posttest. **Study III** extends those findings using a sample of trainee psychotherapists. Also in the target population, both trainings are found to be effective in the one-week posttest. In addition to that, the multimodal training shows effects for unimodal emotion recognition accuracy for audio-only and video-only stimuli in the one-year follow-up towards the end of psychotherapy education. This indicates that standardized computerized emotion recognition accuracy training can be used as a tool for improving trainee psychotherapists’ emotion recognition accuracy, even though additional interventions might be needed for securing long-term success for all facets of emotion recognition. Future research should explore the practical impact of trainee psychotherapists’ emotion recognition accuracy and the training of this ability.

The findings of this thesis are, on the one hand, surprising, in so far as psychotherapy education likely does not lead to improvements in trainee psychotherapists’ emotion recognition accuracy. On the other hand, they are encouraging, because they suggest that this ability can be trained with relatively simple and resource-efficient methods. Emotion recognition accuracy training could become part of standard or individualized psychotherapy training, alongside the training of other relevant verbal and non-verbal socio-emotional and interpersonal psychotherapist competencies.
Sammanfattning på svenska

Forskare, praktiker och lagstiftare är överens om att det är viktigt att förstå vilka typer av psykoterapeutiska behandlingar som leder till varaktiga positiva förändringar i patienters välbefinnande, hur dessa behandlingar kan administreras på effektiva sätt och hur det kan fastställas vilka patienter som skulle dra nytta av vilken behandling. Under de senaste åren har det också funnits ett ökande intresse för dem som utövar psykoterapi; särskilt för de socio-emotionella och interpersonella egenskaper och kompetenser som psykoterapeuter bör besitta för att kunna erbjuda behandling av hög kvalitet för olika patienter, oavsett vilken psykoterapimodul som används. Denna avhandling studerar en sådan viktig kompetens hos psykoterapeuter, nämligen förmågan att känna igen icke-verbala känslouttryck hos andra. Psykoterapeuter behöver kunna hjälpa patienter att uppleva, förstå och uttrycka sina känslor, och i detta sammanhang är det avgörande att de själva har goda socio-emotionella kompetenser, som exempelvis god emotionsigenkänningsförmåga. Trots detta finns det förvånansvärt lite forskning om psykoterapeuters förmåga att känna igen icke-verbala emotionsuttryck och om hur de kan stödjas i att förbättra denna förmåga under sin utbildning.

**Studie I** utforskar blivande psykoterapeuters förmåga att känna igen andras icke-verbala känslouttryck i början och i slutet av den teoretiska och praktiska psykoterapiutbildningen, och jämför dessa terapeuter med en kontrollgrupp av grundutbildningsstudenter. Resultaten visar att blivande psykoterapeuter i början av sin utbildning visar överlägsen förmåga att känna igen multimodala (audio, video, audio-video) känslouttryck och mikrouttryck (<200ms) jämfört med kontrollgruppen. Detta tyder på att de som väljer att bli psykoterapeuter redan initialt kan ha en högre förmåga för icke-verbal emotionsigenkänning, eller att de har utvecklat denna förmåga tidigt under sina studier. Men efter ett och ett halvt års utbildning förbättrar deras emotionsigenkänningsförmåga för multimodala uttryck och mikrouttryck inte signifikant mer än kontrollgruppers förmåga. Detta tyder på att standardutbildning i psykoterapi (psykodynamisk och kognitiv beteendeterapi) inte automatiskt leder till förbättrad emotionsigenkänningsförmåga, även om blivande psykoterapeuter lär sig att bedriva psykoterapi och även behandlar sina första patienter vid universitetskliniken. En alternativ förklaring är att de socio-emotionella kompetenser som utvecklas under utbildningens gång inte fängas av de standardiserade datori-
serade test för emotionsigenkännande som används i denna studie. Denna upp-
täckt indikerar ändå att mer explicit träning i icke-verbal emotionsigenkännande
är nödvändig under psykoterapiutbildningar. Studie II undersöker sedan två
nyutvecklade standardiserade datoriserade träningsprogram för emotions-
igenkännande, ett för multimodal emotionsigenkännande och ett för förmågan att
cänna igen mikroutryck. Träningsprogrammen utvärderas i en grupp grund-
utfinningsstudenter och använder sig av en mixed-design. Träningsprogram-
men jämförs med varandra och med en aktiv kontrollträning. Båda tränings-
programmen visar sig signifikant förbättra deltagarnas emotionsigenkän-
ningsförmåga i ett posttest en vecka senare. Studie III utvidgar dessa resultat
genom att inkludera blivande psykoterapeuter som deltagare. Även i denna
målgrupp visar sig båda träningsprogrammen vara effektiva i ett posttest en
vecka senare. Dessutom visar två delmoment i det multimodala träningspro-
grammet effekter av unimodal emotionsigenkänande av endast-ljud och end-
ast-videostimuli i en ettårs-follow-up mot slutet av psykoterapiutbildningen.
Detta tyder på att standardiserade datoriserade träningsprogram för emotions-
igenkännande kan användas som ett verktyg för att förbättra blivande psykote-
rapeuters emotionsigenkänningsförmåga, även om ytterligare insatser kan be-
hövas för att säkra långsiktiga resultat för alla aspekter av icke-verbal emo-
tionsigenkänning. Framtida studier bör utforska praktisk påverkan i terapisitu-
ationen av blivande psykoterapeuters emotionsigenkänningsförmåga och hur
träning av denna förmåga påverkar terapeut-patient-relationen.

Resultaten i denna avhandling är å ena sidan överraskande, eftersom psy-
koterapiutbildning inte verkar leda till förbättringar i blivande psykoterapeu-
ters emotionsigenkänningsförmåga. Å andra sidan är de uppmuntrande, eft-
ersom de antyder att denna förmåga kan tränas med relativt enkla och resur-
seffektiva metoder. Samtidigt indikerar resultaten att träning i emotions-
igenkännande bör bli en del av standard- eller individualiserad psykoterapiutbildning, tillsammans med träning av andra relevanta verbala och icke-verbala socio-emotionella och interpersonella psykoterapeutiska
kompetenser.
List of publications

Döllinger, L., Letellier, I., Högman, L., Laukka, P., Fischer, H., & Hau, S. (2023). Trainee psychotherapists’ emotion recognition accuracy during 1.5 years of psychotherapy education compared to a control group: No improvement after psychotherapy training. (accepted by PeerJ, 14th of September 2023)


## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>CBT</td>
<td>Cognitive Behavioral Therapy</td>
</tr>
<tr>
<td>ERA</td>
<td>Emotion Recognition Accuracy</td>
</tr>
<tr>
<td>ERAM</td>
<td>Emotion Recognition Assessment in Multiple Modalities test</td>
</tr>
<tr>
<td>GEMEP</td>
<td>Geneva Multimodal Emotion Portrayals Core Set</td>
</tr>
<tr>
<td>MICRO</td>
<td>Micro expression recognition task</td>
</tr>
<tr>
<td>PDT</td>
<td>Psychodynamic Therapy</td>
</tr>
<tr>
<td>PECT</td>
<td>Patient Emotion Cue Test</td>
</tr>
</tbody>
</table>
Acknowledgements

Stephan Hau, you were more than a supervisor, you are a real mentor to me. You have opened up so many doors for me, always having such faith in me, and being so generous with your encouragement to pursue new endeavours, even though you knew that they sometimes would lead me far away from my dissertation project. It seems to me that, for you, it was not only about me producing data and articles, but about me finding my own way and growing as a scientist, clinician and person. I admire your creativity and your willingness to forego the easy way in favour of seeking new challenges in your research, as well as your integrity as a person. Thank you for your endless patience with my neurotic nature, for your unwavering support and trust, and for all the last-minute, late-night emails that you answered during this long process! I hope to continue working with you for a long time.

Thank you to the whole Wallenberg research team! Håkan Fischer, thank you for being my second supervisor and for your encouragement and positive attitude throughout this time. You always had the most cutting-edge solutions and suggestions for improving the project and it was such a relief seeing you radiate this boundless optimism for the project and when giving career advice. Isabelle Letellier, thank you for trusting me with your participants and with your data. When I was new in Sweden, you took me under your wing and shared your experience with me. You were both a mentor and a friend and I hope to have many more conversations with you about surviving in academia, climate issues and life. Lennart Högman, without you there would have been no project and no thesis, because there would not have been any data collection. You were the master of the lab and technology and I can’t thank you enough for being such a support in the beginning of the project. Petri Laukka, you were somewhat of an informal supervisor to me, and I’m so glad that I always could turn to you when it came to research methods, test design, statistics and all the things emotion recognition. I can’t believe how fast you managed to answer all my questions and to have discussions on such a deep level, considering that you are part of a gazillion research projects. Thank you so much for that! Tanja Bänziger, you were a mentor to me early on and I was so amazed by all your knowledge in emotion recognition research and beyond and by your passionate way of sharing it with me. I would have loved to work with you more. Irena Makower, it was such a joy to have you on
board in the research team, not only because you brought in fresh wind as the only CBT therapist, but also because you have this vast pool of ideas and expertise. Diana Persson, you stepped up for the parts of the project that I had less experience in, like the MRI sub-study and data collection, and while I was on parental leave. Thank you so much that I always could count on you! And Kristoffer Månsson and Amir Manzouri, thank you for contributing all your knowledge to the MRI project. Thank you also to my other co-authors and colleagues that I have worked with in other projects throughout the years!

Thank you to all the research assistants and students who helped us with data collection, digitalisation and data management, especially to Carlos Tirado Aldana, Malina Szychowska, Patrick Schuster, Emma Neal, Alica Klein, Felicitas Schmidt, Sophie Albrecht, Tova Magnusson, Amanda Spejare, Hanna Nilsson, Hellen Verhoeven, Fanny Antonsson, and Linnéa Engström. In case I forgot to mention any of you, I’m sincerely sorry and thank you as well! Without you, this massive project would not have been manageable!

My sincerest thank you also goes to all the former and current colleagues at Stockholm University’s Psychological Clinic, specifically to Philip Stenström, Anastasia Oysiannikova, Jenny Bourelius, Karin Pålsson, Anna-Clara Hellstadius, Aina Lindgren, Thomas Lindgren, Alexandra Billinghurst, and Anna Mautner. It was a very demanding and time-consuming task to collect the outcome and process data at the clinic, and it was you who made this research possible!

I also would like to express my appreciation to the university and the funding agencies that supported my research and the articles presented in this thesis. Thank you to Stockholm University, to the Marcus and Amalia Wallenberg Foundation and to Forte and the Marie Skłodowska-Curie Actions from the EU commission.

Furthermore, I would like to express my utmost gratitude to the trainee psychotherapists, clients and undergraduate students that chose to be part of our research. This was not an easy study and that so many of you came back to the lab over and over again is the basis for this thesis. Thank you!

Thank you to the two reviewers of an earlier draft of this thesis – Ann-Charlotte Smedler and Yvonne Brehmer – and to the two half-time reviewers – Ewa Mörtberg and Andreas Gerhardsson – for your incredibly thoughtful and valuable comments for improving the manuscript! Thank you also the reviewers and publishers of the articles presented in this thesis.
A sincere thank you to the committee for my PhD defence – Uta Sailer, Håkan Johansson, Lilianne Eninger, and Gebrenegus Ghilagaber Yebio – for agreeing to evaluate my work, and to the chairman of the defence, Björn Philips. I’m looking forward to seeing you at the day of the defence! And a special thank you in advance to my opponent Patrick Luyten. I’m looking forward to meeting you and to discussing my work with you!

Apart from my supervisors and all the people that contributed to the studies this thesis is based upon and to the defence, there were so many other people that helped and inspired me along the way.

Pehr Granqvist, you were the person that made this all possible for me. When I was a master’s student from Germany looking for an internship, you didn’t hesitate to offer me a job in a fascinating study. You took a chance on me and since then we have published two articles together; I hope there will be more to come in the future. You are one of the brightest and most passionate people I have ever met and I thank you for all the stimulating conversations, with you personally, and in the attachment club, that kept the flame in me alive during some uninspired times during my PhD journey. Thank you! Also thank you to my fellow clubbers, especially to Mårten Hammarlund, Joel Gruneau Brulin, Aaron Cherniak, Tommie Forslund, Anna Wallenberg, and Freja Isohanni for so many fascinating discussions.

Alexandra Billinghurst, it is immeasurable what you did for me. You took in a stranger, so that I could enter the psychotherapy education here in Sweden. And you did it out of the kindness of your heart, as well as for the greater cause of helping create research within psychodynamic therapy and psychoanalysis. I’m amazed by how you look out for the clients at our clinic and for the trainee psychotherapists of the clinical psychology program, as well as by the deep knowledge that you share so selflessly. You are a role model for so many of us and I hope to live up to your example one day and to show the same kindness to new generations of PDT therapists. Thank you!

Thank you for stimulating conversations and collegiality to the current and former colleagues at the department of psychology, especially to the colleagues at the clinical division. Among others, many thanks to Stephan Hau, Herman Daniels, Camilla von Below, Björn Philips, Karin Lindqvist, Jacob Mechler, Robert Johansson, Nora Choque Olsson, Jan Bergström, Alexandra Billinghurst, Christer Ekelund, Gunnar Berggren, Peter Lillengren, Charlotta Björklind, Bo Vinnars, and Per Carlbring. Thank you also to the heads of the different study programs that have given me the chance to teach a lot and early on during my career and that trusted me with their courses and students, especially Lilianne Eninger, Laura Ferrer Wreder, Jan Bergström and Gunnel Jacobsson.
Furthermore, my heartfelt thank you to the many colleagues working in administration, IT, HR and other capacities at the department that have made my life as PhD student and teacher so much easier! Thank you also to all the people that make a difference for PhD students at the department and the university, like the PhD council.

My closest friends and colleagues at the department, thank you! First and foremost, my good friend Diana Persson, you have been on this road with me since the very beginning. Without you, I would still be trying to figure out most of the going-ons at the department and in the lab. You were always there for me, at work and in life, and I could not have wished for a better partner in crime than you! Thank you for sticking by me and for all your support! Hellen Vergoossen, thank you for all the interesting conversations, your sense of humour and for your friendship. Anna Wallenberg, thank you for your light-heartedness and good conversations. Marta Zakrzewska and Joanna Lindström, thank you for being such wonderful companions on the road. Karin Lindqvist and Jacob Mechler, thank you for fighting the good fight of bringing science into PDT. Carlos Tirado Aldana and Louise Bergman, thank you for your integrity and your black sense of humour. Thank you to Katarina Bendtz, Azade Azad, Miriam Zehnter (and Diana, Hellen and Marta) for the Women in Psychology network. Thank you to Johanna Schwarz and Loretta Platts for being great role models. Thank you also to Mårten Hammarlund, Joel Gruneau Brulin, Andreas Gerhardsson, Alexander Miloff, and the many other PhD students that I was lucky enough to get to know during the years.

I’m gratefully looking back to my psychology studies in Bamberg, Germany. For the first time in my life, studying felt meaningful and I felt like I’d found a purpose. I would like to thank my teachers at the Otto-Friedrich University, in particular Prof. Michael Hock, who was the supervisor for my bachelor’s and master’s theses. Thank you for putting me on the right track in learning how to use R and for kindling the idea in my mind that I actually might be able to become a researcher. I would also like to thank my learning group in Bamberg. When I arrived, I had no idea how to study for an exam or how to structure myself at all. You have taught me all those skills and I’m so glad you took me under your wings. Thank you, Jana, Karo, Manu, Anna and Krissi! Thank you also to the division of personality psychology for my first jobs as research assistant.

Thank you to my closest friends outside of academia (yes, there is a life outside of academia!), especially Mitschka, Jana, Karo, Franzi, Kim, Britta, Klaudia, Nadine, Natalie and Volkan for sharing laughs and tears with me throughout the years.
Thank you to the psychology and psychotherapy students that I have met during my years as a teacher at the various programs at the university. The saying that teaching is the best way of learning is absolutely true. It is a joy seeing you become such competent and dedicated psychotherapists! Thank you also to the many teachers I have had throughout my life, especially at the psychotherapy program.

Thank you to my therapists, the one who encouraged me to go my own way and to study psychology, and the one who held my hand through the rough patches in the end of this journey. Thank you also to my clinical supervisors for their wisdom and dedication!

Thank you to my clients, who have moved me in just as many ways as I hopefully have affected them.

And lastly, I would like to thank my family for all that you did for me and for simply being there, Mom, Dad, Lana, Dominik, Roland, Sabine and River.
# Contents

Abstract .......................................................................................................................... 1
Sammanfattning på svenska .......................................................................................... 3
List of publications ........................................................................................................... 5
Abbreviations .................................................................................................................. 6
Acknowledgements ........................................................................................................ 7
Contents ........................................................................................................................... 13
Introduction ..................................................................................................................... 15
  Therapist characteristics and competencies .............................................................. 16
  Emotion recognition accuracy .................................................................................. 18
Emotion recognition accuracy in psychotherapy ......................................................... 20
  Psychotherapists’ emotion recognition accuracy ...................................................... 22
  Emotion recognition accuracy training for psychotherapists .................................... 24
  Emotion recognition accuracy facets of interest ....................................................... 26
    Multimodal emotion recognition accuracy ............................................................ 26
    Micro expression emotion recognition accuracy .................................................. 26
Overview of the thesis .................................................................................................... 28
  Aims and background ............................................................................................... 28
  Participants .................................................................................................................. 29
  Emotion recognition accuracy measures .................................................................. 29
  Emotion recognition accuracy trainings .................................................................... 32
  Ethical considerations ............................................................................................... 35
  Research projects and funding .................................................................................. 36
  Overview of the studies ............................................................................................. 37
Study I .............................................................................................................................. 38
  Aims and background ............................................................................................... 38
  Method ....................................................................................................................... 39
  Results ......................................................................................................................... 39
  Discussion and conclusion ......................................................................................... 41
Study II ............................................................................................................................. 43

13
Aims and background ........................................................................................................... 43
Method ................................................................................................................................. 43
Results ................................................................................................................................. 44
Discussion and conclusion ................................................................................................. 45

Study III ............................................................................................................................... 47
Aims and background ......................................................................................................... 47
Method ................................................................................................................................. 47
Results ................................................................................................................................. 48
Discussion and conclusion ................................................................................................. 51

Discussion ............................................................................................................................ 53
Summary of results .............................................................................................................. 53
  Similarities and differences of the results ........................................................................ 54
Discussion of the results and implications ....................................................................... 56
  Trainee psychotherapists' ERA is higher than undergraduate students' ...................... 56
  No changes in ERA from beginning to end of psychotherapy education .................... 58
  Considerations about the undergraduate students' ERA ............................................ 60
  Successfully training ERA ............................................................................................ 61
Methodological considerations ......................................................................................... 63
  Study design and statistical methods ............................................................................ 63
  Reliability issues .............................................................................................................. 65
  Specifics of the tasks and trainings ............................................................................... 66
  Psychotherapeutic approach .......................................................................................... 67
Summary of strengths and limitations .............................................................................. 68
Open questions and future directions .............................................................................. 69
Conclusion and clinical relevance ...................................................................................... 73

References ............................................................................................................................ 74
Introduction

The World Health Organization’s (WHO, 2022) mental health report suggests that about 970 million people worldwide lived with a mental disorder in 2019, not including the millions living with alcohol or drug use disorders. In a large survey by the Public Health Agency of Sweden (Folkhälsomyndigheten, 2023), 11% of women and 5% of men living in Sweden in 2022 reported severe mental health difficulties, and 41% of the women and 30% of the men reported mild difficulties (e.g., mild anxiety and stress). These are huge numbers of people suffering from oftentimes treatable mental disorders and difficulties. The most prevalent disorders, according to the WHO (2022) were depressive and anxiety disorders, both of which we know respond well to psychological treatments (see, e.g., Munder et al., 2019; Wampold & Imel, 2015). Although there is still debate about the effect sizes for psychotherapeutic and psychopharmacological treatments (see, e.g., Leichsenring et al., 2022), research and the experience of patients and therapists alike provides good evidence for the positive effects of psychotherapy (e.g., cognitive behavioral therapy, CBT, and psychodynamic therapy, PDT) on mental health. Still, it is less clear, how or why positive changes in psychotherapy take place.

There is a wealth of research trying to define which factors determine whether a psychotherapy will be successful. Meta-analyses and reviews (see, e.g., Norcross & Lambert, 2019) suggest that there are several variables that contribute to psychotherapy outcome. Beyond factors that are happening outside of therapy and expectancy/placebo effects, the contents of psychotherapy and how psychotherapy is facilitated have been extensively studied. When investigating which therapeutic factors are contributing to explaining psychotherapy results, it is generally found that the patient’s contribution (e.g., motivation for treatment, severity of problem) in explaining variance in outcomes is about 30%, the therapeutic relationship’s contribution is about 15%, the treatment method accounts for about 10% of the variance, the individual psychotherapist for about 7% and other factors for about 3%. Notably, there is also a large proportion of unexplained variance (35%). These numbers vary somewhat from study to study. Further, it is important to keep in mind that all of those factors work together in determining the outcome of a psychotherapy and that research in all of those areas, and about their interplay, is needed (Norcross & Lambert, 2019).
The present thesis aims to add to the research about individual psychotherapists’ contributions to psychotherapy. This line of research is, in my view, particularly relevant, because it is an intuitive point of action in psychotherapy educations. In psychotherapy programs, we mostly work with trying to support trainee psychotherapists in learning how to provide the best possible treatments for their patients. Thus, psychotherapy education research tries to work out exactly how psychotherapists can influence the course of a therapy. This includes specific psychotherapeutic techniques and interventions relevant for the specific treatment approach, as well as individual therapist factors and how the therapist contributes to establishing a positive therapeutic relationship with a variety of patients. Another important question in the field is how trainee psychotherapists can be supported in learning both techniques and important therapist characteristics and competencies, that are beneficial for psychotherapy outcome and process (e.g., relationship or alliance-building and maintaining).

In the present thesis, I set out to understand a specific socio-emotional therapist competency, namely the ability to recognize non-verbal emotional expressions in others, and how this ability changes in response to psychotherapy education. Further, I wanted to find out whether emotion recognition accuracy (ERA) could be trained with standardized computerized methods as part of the psychotherapy education.

Therapist characteristics and competencies

In recent years, there has been considerable interest in the individual psychotherapist’s contribution to psychotherapy outcome and process (see, e.g., Norcross & Lambert, 2019; Orlinsky & Rønnestad 2005, Wampold et al., 2017). This might be in part due to a general interest in factors that could be important in all forms of psychotherapy irrespective of approach, so called common factors (see, e.g., Wampold, 2015), and the debate surrounding the “dodo bird verdict”, i.e., the idea that most therapy approaches lead to relatively similar results (Luborsky et al., 1975, Luborsky et al., 2002). Empirically, the findings for the question whether different therapy approaches yield different or similar results are somewhat inconclusive and suggest that the answer to this question is likely more complex (Beutler, 2002; Chambless, 2002; Cuijpers, 2016). It certainly is beneficial to try to determine which kinds of treatment might lead to which results in which patients and to adopt a more personalized approach to psychotherapy (Barber, & Solomonov, 2019; Roth, & Fonagy, 2005). Further, it should also be investigated who (i.e., which therapist) works for whom and why (Constantino et al., 2017). Nowadays, there is increasing interest in the question of how the individual therapist influences therapy results and how therapists can be supported in developing relevant competencies (e.g., Castonguay, & Hill, 2017; Norcross, & Lambert, 2019).
Several meta-analyses highlight the role that psychotherapists themselves play in explaining the success (or failure) of a psychotherapy. Depending on the meta-analysis or review, the contribution of the individual therapist, the so-called therapist effect, has been found to explain between three to eight percent of the variance in therapy outcomes (Baldwin & Imel, 2013; Heinonen & Nissen-Lie, 2020; Johns et al., 2019; Norcross & Lambert 2006; 2019; Wampold & Imel 2015). Some studies even report effects of up to 29% and suggest that therapist effects are particularly strong in therapies with more challenging patients (Johns et al., 2019; see also Holmqvist, 2020). The search for characteristics of “supershinks” (Ricks, 1974), therapists that produce exceptional results with their patients, has been long. Research shows that some psychotherapists are much more effective than others. For example, Okiishi et al. (2003) found that the patients of the most effective psychotherapists in their study showed ten times greater change rates than the average improvement. The patients of the least effective psychotherapists in the study even deteriorated. Brown et al. (2005), e.g., found that the patients of the most effective therapists in their sample had three times greater symptom reduction compared to patients of the other therapists. Therapist effects have been found to be stable, meaning that the most exceptional therapists even performed that way with future patients (Kraus et al., 2016), and they even play a role for treatment length, dropout rates, early change and sudden gains (see Delgadillo et al., 2022 for a review). Examples for established therapist effects are the ability to form an alliance (i.e., being able to establish agreement about the goals and tasks of therapy and the ability to form a bond between patient and therapist, see Bordin, 1979) across a range of patients (e.g., Flückinger et al, 2018), a variety of facilitative interpersonal skills (including empathy, warmth, verbal fluency, appropriate expression of emotion, persuasiveness, communication of hope, the ability to create a bond with patients, and being able to focus on the patients’ problems, see, e.g., Anderson et al., 2009), professional self-doubt (in the sense that the more effective therapists were the ones with more self-doubt, see, e.g., Nissen-Lie et al., 2013), or engaging in deliberate practice (i.e., the targeted training of specific aspects of the individual performance, Ericsson et al., 1993; see, e.g., Chow et al., 2015) (for reviews, see, e.g., Holmqvist, 2020; Wampold et al., 2017). Therapist effects have even been found in trainee psychotherapists at the graduate level (see, e.g., Edmondstone et al., 2023), suggesting that they are important right from the beginning of psychotherapists’ careers.
Emotion recognition accuracy

Non-verbal ERA can be broadly defined as the ability to recognize other people’s emotional states without verbal input, i.e., solely based on non-verbal behavior, like facial expressions, bodily posture and tone of voice (prosody). Expressing emotions via non-verbal channels has interpersonal benefits, e.g., it allows for communicating inner states and by that to influence other people’s behavior (e.g., to signal for help in dangerous situations) and inner states (e.g., evoking fear through expressing anger). To recognize emotional expressions in others in turn allows to draw conclusions about other people’s inner states and influences own behavior (e.g., to provide comfort to a sad friend). Being able to recognize how an interaction partner might feel without them having to verbally express their emotional state has been found to be positively associated with mental health and well-being, psycho-social functioning, social and work-related competencies, and other positive traits (see, e.g., Bänziger, 2016; Hall et al., 2009; Schlegel et al., 2013). Healthy adults are relatively good at accurately recognizing other people’s emotions compared to clinical populations, like individuals high up on the autism spectrum (Uljar-evic & Hamilton, 2013; Yeung, 2022), with schizophrenia (Kohler et al., 2010), depression (Dalili et al., 2015), antisocial traits (Marsh & Blair, 2008) or other mental health issues. Even in healthy individuals, it is found that low ERA is associated with a range of maladaptive traits and interpersonal problems (see, e.g., Hall et al., 2009; Schlegel, Fontaine, et al., 2017).

Generally, individuals vary in their ability to recognize emotional expressions in others (Bänziger, 2016). Some variability in ERA can be explained by methodological features of ERA tests. There are, e.g., measures for different ERA facets, like expressions of emotions below and above a conscious detection threshold. Emotional expressions are generally considered micro expressions under a threshold of 200ms, even though more recent research suggests that expressions with a threshold of below 500ms could still be classified as micro expressions (see Matsumoto & Hwang, 2018). They are standing in contrast to spontaneous non-concealed expressions of “normal” duration, sometimes called macro emotions. Micro expressions are generally harder to recognize than expressions with longer durations (see, e.g., Frank & Svetieva, 2015). Emotional expressions are expressed, perceived and recognized via different channels of expression/modalities (e.g., face, body, voice, gestures, displays) and people can vary in their ERA for different modalities or emotions. Generally, unimodal displays of emotion are more difficult to recognize than the recognition of combined audio-visual displays of emotion, and auditory displays tend to have lower accuracy rates than visual displays (Cortes et al., 2021; Laukka et al., 2021). Emotion recognition research also varies in regards to which emotions are included in tasks or which theory of emotion is followed.
Most ERA research stands in the tradition of the *theory of basic emotions*, that goes back to Charles Darwin’s (1872) seminal book *The expression of emotion in man and animals*, but there were also other important influences. Darwin suggested that emotional expressions are similar in humans and non-human animals and that they serve a non-verbal communicative function. In accordance with his theory of evolution, he was underlining the adaptive function of emotional expressions and their recognition. Silvan Tomkins and later Paul Ekman built on Darwin’s work and formulated the theory of basic emotions that has stimulated much research in the field during the last decades. Ekman (see, e.g., Ekman & Cordaro, 2011) is defining basic emotions as congenital automatic responses that can be found universally in cultures around the globe, even though the descriptions and attitudes towards those emotions can vary. He also points to the distinctness of basic emotions and their expression, e.g., distinct physiological changes and facial muscle contraction patterns. There is evidence for the universality of the basic emotions *anger, fear, sadness, disgust, surprise* and *happiness* (see, e.g., Elfenbein & Ambady, 2002). Social constructivist theories of emotion argue against universality and biological determination of emotions and emotional expressions, but instead suggest that they are individually and socially constructed, based on past experiences and cultural factors (see, e.g., Barrett, 2017). Other theories suggest that emotions and emotional expressions can include universal as well as culture-specific components and that the individual appraisal of the situational context determines how emotions are interpreted and expressed (see, e.g., Scherer et al., 2011).
Emotion recognition accuracy in psychotherapy

In the present thesis, I investigated trainee psychotherapists’ ERA and how it can be trained with standardized computerized training procedures as part of psychotherapy education. These research questions go back to the idea that ERA could be a therapist factor that is contributing to better therapy results.

Psychotherapists need to be able to identify patients’ emotional expressions accurately to get a better and more empathic understanding of their patients’ inner worlds and to help them understand themselves and their own and other people’s emotions in new ways. They need to, e.g., help patients perceive, experience, label, and reflect about their (often complex) emotions. Patients’ non-verbal emotional expressions can possibly guide psychotherapists in planning and applying fitting interventions (e.g., emotion exposure, emotion regulation, adequate expression), in diagnosing and in risk assessment (see, e.g., Ekman & Friesen, 1969). In addition, being able to read patients’ emotional expressions (without the patients having to verbally express how they are feeling), and to successfully work with that knowledge, might contribute to establishing a good working relationship or alliance with patients (Bennett-Levy & Thwaites, 2007; Hutchison & Gerstein, 2012) and help facilitating corrective emotional experiences for patients. This leads to the idea that ERA might not only be a therapist factor for outcome but also a contributing factor to the therapeutic relationship.

Another reason for why psychotherapists’ ERA might be important is that many patients display deficits in emotion recognition, e.g., individuals with autism spectrum disorder (Uljarevic & Hamilton, 2013; Yeung, 2022), schizophrenia (Kohler et al., 2010), depression (Dalili et al., 2015), antisocial traits (Marsh & Blair, 2008), and possibly borderline personality disorder (Domes et al., 2009), even though the research on this latter population is mixed (see van Heel et al., 2019). These and other clinical populations might be in need of exploring their own and other people’s emotions together with a therapist. To facilitate this emotion focus, the psychotherapist likely needs to possess good ERA themselves. Further, many specialized psychotherapy approaches build on affect theory (see, e.g., Bergsten et al., 2015). Those approaches focus on the joint exploration of emotions in session, which includes experiencing, labeling and regulating emotions, among other aspects. Examples are *Emotion*
Focused Therapy (Greenberg, 2015), Intensive Short-Term Dynamic Psychotherapy (Coughlin, 1996; Davanloo, 2001), Affect Phobia Therapy (McCollough et al., 2003), Accelerated Experiential-Dynamic Psychotherapy (Fosha & Yeung, 2006), Dialectical Behavior Therapy (Linehan, 2015), Unified Protocol (Barlow et al., 2017), or Acceptance and Commitment Therapy (Hayes, & Lillis, 2012). However, most psychotherapeutic schools consider working with emotions an important part of treatment (see, e.g., Ehrenreich et al., 2007; Greenberg & Safran, 1989; Greenberg et al., 2019; Hofman, 2015; Hutchison & Gerstein, 2012), which indicates it to be a common factor in psychotherapy.

Whether ERA really contributes to psychotherapy outcome and process is still understudied. In the, to my knowledge, first study that systematically investigated this question, Abargil and Tishby (2021) used a standardized computerized task for ERA to assess, in a more objective manner, whether psychotherapist empathy was related to therapy outcome and process. In this study, the authors used ERA as a proxy for empathy. Whilst ERA is not the same as empathy, the operationalization of empathy via ERA tasks lies close to hand, as ERA likely covers at least parts of the empathy construct. In theory, ERA is often equated with cognitive, perspective-taking components of empathy (i.e., being able to read another person’s emotion), in contrast to affective/emotional empathy (i.e., being able to feel another person’s emotion; see, e.g., Martingano & Konrath, 2022). Empirically, the correlational patterns are more complex (see Besel & Yuille, 2010). Emotion recognition accuracy is also sometimes seen as a prerequisite for empathy (Besel & Yuille, 2010). Still, it has to be noted that more recent research questions whether ERA and empathy have strong co-variation. Flykt et al. (2023) found that ERA was only a weak predictor of empathic accuracy, suggesting that they might be separate competencies. Possibly, empathy and ERA, whilst related might belong to a broader construct, like interpersonal accuracy (Hall et al., 2016).

However, in their study, Abargil and Tishby (2021) found that psychotherapists’ ERA was a moderator of the patients’ change in four outcome and four process variables. The patients of therapists with high ERA reported less severity of their three main target complaints over time, and that they became better at regulating their emotions, more accepting of their emotional responses and that they had more emotional clarity over time (outcome variables). Further, they reported that they had a stronger secure attachment to the therapist over time, and there was a trend for less avoidant attachment to therapist over time. Psychotherapists with high ERA were able to establish a stronger working alliance from session to session than the ones with less ERA, and the patients of therapists with low ERA reported feeling more tense and upset during the sessions, which the authors interpreted as possible alliance ruptures and non-resolutions (process variables).
Apart from the mentioned study, research about psychotherapists’ emotional competencies in relation to therapy outcome and process mostly investigated psychotherapists’ empathy levels. Meta-analytic research finds that psychotherapist’s empathy is an at least moderately strong moderator of outcome, although this relationship varies depending on which empathy measures are used (e.g., client ratings, observer ratings, self-reports, or congruence scores of these measures) (Elliott et al., 2011; Watson et al., 2014). In a meta-analysis, it was found that the empathy–outcome relationship was stronger for less experienced psychotherapists (e.g., Elliott et al., 2011). Empathy, like other therapist characteristics, seems to be most impactful for less experienced therapists. Further, there are also studies that found a strong positive association between therapist empathy and therapeutic alliance (Nienhuis et al., 2018) and a negative relationship between therapist empathy and patients’ ratings of attachment insecurity (Watson et al., 2014; Watson et al., 2020). It is reasonable to believe, that ERA follows a similar pattern. However, this assumption needs to be further tested in future studies about psychotherapists’ ERA and therapy outcome and process. In the course of my PhD studies, I have also collected therapy outcome and alliance data for the patients of the trainee psychotherapists that were participating in Studies I and III in this thesis. These results are pending.

Psychotherapists’ emotion recognition accuracy

There are only a few systematic studies about psychotherapists’ and trainee psychotherapists’ ERA. Some investigate the question whether this population is better at detecting non-verbal emotional expressions in others than other populations. Others engage with the question whether clinical experience or psychotherapy education might lead to improvements in ERA. Theoretically, it makes sense to assume that psychotherapists that have long experience with working with patients’ and their emotions might possess better ERA, since the frequent and intensive work with patients’ emotions might stimulate this ability. One could say that clinical experience could improve psychotherapists’ ERA competencies. However, it is still not entirely clear whether psychotherapists actually are more successful in reading other peoples’ emotions than other populations. Higher ERA or higher interest in other people’s emotions might even influence young psychotherapists’ decision to choose this profession. One could say that higher ERA – as a personal characteristic – could lead to seeking those kinds of career paths.

Machado et al. (1999) investigated how well experienced psychotherapists ($n = 36$) and undergraduate clinical psychology students ($n = 36$) who intended to become psychotherapists were able to recognize patients’ verbal and non-verbal emotional expressions. The ERA measures in this study were a video
recording of a psychotherapy session (naturalistic condition), an altered version of the video in which the verbal content was filtered out (only non-verbal ERA) and, lastly, a verbatim transcript of the recorded session (only verbal ERA). The therapy session was rated by experts that came to a consensus on which primary emotions were displayed by the patient (love, joy, surprise, anger, sadness, and fear). Both groups completed all three conditions. The study found that the experienced psychotherapists, overall, were better at recognizing patient’s emotional expressions in all three conditions. The non-verbal ERA condition showed the strongest contrast between psychotherapists and undergraduate students. The undergraduate students apparently tended to rely more on the combined input of verbal and non-verbal expressions. The authors conclude that the experienced psychotherapists made more use of the non-verbal input, whereas the undergraduate students focused on the verbal content. The difference between experienced practitioners and students suggests that non-verbal ERA could increase in the course of clinical experience and education. However, since this was a cross-sectional naturalistic study, more, and longitudinal research is needed. Pauza et al. (2010) explored whether trainee psychotherapists \((n = 121)\) in the beginning of their education showed higher ERA rates than trainee coaches \((n = 17)\), a normal population sample \((n = 250)\) and a clinical sample of patients with anxiety disorders \((n = 37)\). In contrast to Machado et al., Pauza et al. used a standardized ERA measure that included pictures of faces displaying the basic emotions fear, joy, surprise, disgust, sadness and anger. The results suggest that the trainee psychotherapists were superior to the other three groups and that there were no differences between the included therapy approaches (CBT, PDT, PDT for children, psychodrama). The authors discuss that individuals with high interpersonal communication competencies might be more interested in pursuing a psychotherapy education. These studies support the idea that (trainee) psychotherapists possess superior ERA and that clinical experience and education might be associated with higher ERA.

Contrary to those findings, Hassenstab et al. (2007) did not find a difference between practicing psychotherapists \((n = 19)\) and matched controls \((n = 19)\) regarding the ability to recognize the basic emotions happiness, sadness, anger, fear, disgust, surprise, and neutral facial expressions using a standardized ERA task including pictures of facial expressions. Similarly, Hutchison and Gerstein (2012) did not find a difference between counseling trainees \((n = 54)\) and undergraduate students \((n = 54)\) in a recognition accuracy task using pictures of the basic emotions anger, contempt, disgust, fear, happiness, sadness, and surprise. These results speak against the idea that (trainee) psychotherapists possess superior ERA.

There are some studies about other mental health professionals’ ERA. Two studies, e.g., found that clinical experience and education might be associated with increased ERA. One of them investigated psychiatry residents during
three years of residency (Arango de Montis et al., 2014) and the other compared mental health professionals and nursing students (Minardi, 2013). Another study (Ragsdale et al., 2016) could not find any significant difference between medical faculty and medical students.

To summarize, the research about psychotherapists’ and trainee psychotherapists’ ERA is inconclusive and more research in this field is needed to find answers for the questions whether (trainee) psychotherapists’ have superior ERA before they start psychotherapy training, and, possibly even more importantly, whether psychotherapy education and experience can lead to improvements in ERA. Study I of this thesis aims at contributing to the investigation of these questions by following a sample of trainee psychotherapists from the beginning to the end of their specialized (PDT or CBT) psychotherapy training and comparing them to a sample of undergraduate students.

Emotion recognition accuracy training for psychotherapists

Assuming that psychotherapists’ ERA indeed is associated with positive therapy outcome and process, and given the unclear picture about (trainee) psychotherapists’ ERA and whether this ability improves due to clinical experience and education, it is relevant to investigate whether (trainee) psychotherapists’ ERA could be improved with interventions. We know from emotion psychology research that ERA can be successfully trained with the help of standardized (computerized) training programs (for overviews, see, e.g., Rebeschini et al., 2019; Schlegel, Vicaria, et al., 2017; or Study II of this thesis). Standardized computerized ERA training during psychotherapy education might be an effective and resource-efficient way to help trainee psychotherapists to improve their ERA. However, there is currently only limited research conducted about the systematic training of (trainee) psychotherapists’ ERA.

So far, I know of only two randomized controlled studies about ERA training for psychotherapists that were also evaluated in samples of (trainee) psychotherapists. Johnsen (2018) trained one half of a sample of fifty-five clinical psychologists, counsellors, and psychotherapists with the help of Ekman’s Subtle Expression Training Tool (Paul Ekman Group, 2022), a training for detecting more subtle emotional expressions. The other half did not receive any training. The training group performed significantly better at recognizing a patient’s non-verbal emotional expressions in a video-filmed session two weeks after the training. Curtis (2021) used Ekman’s Micro Expression Training Tool (Paul Ekman Group, 2022) to train a sample of twenty-three graduate level counseling students and thirty-two undergraduate students in micro expression ERA. They found that the computerized training for micro expression
recognition accuracy led to improvements from pretest to immediate (same day) posttest, in comparison to a control condition in which the participants only watched a video of a therapy session. Specifically, there were significant improvements for the emotions contempt, anger, and fear. They did not find any differences between the graduate counseling students and the undergraduate students when it comes to micro expression ERA. The authors had also hoped that the training would improve deception detection (measured via video clips of people expressing true statement and lies). However, this was not the case.

There were also some studies attempting to train other mental health professionals in ERA. Yu et al. (2016), for example, showed that medical students that underwent a micro expression and subtle expression training became significantly more accurate than their colleagues undergoing a control training. Endres and Laitlaw (2009) found that first-year medical students with high communicative skills became significantly more accurate in detecting facial micro expressions after undergoing a micro expressions training, but not their colleagues with low communicative skills. Two studies that utilized empathy and relational skills or interpersonal process training (Riess et al., 2011; Robbins et al., 1979) found no significant increase in ERA for medical residents, whereas another (Riess et al., 2012) found a significant improvement compared to standard residency training (for a review, see Blanch-Hartigan & Ruben, 2013); however, these trainings did not explicitly target ERA. Another study (Ragsdale et al., 2016) reported an improvement in interpreting facial emotional expressions in medical faculty and students after learning about and training with facial expression features of basic emotions. Blanch-Hartigan (2012) developed an ERA training tailored for health care providers that was found to be effective, however, the training used verbal and non-verbal expressions simultaneously and was evaluated using a sample of undergraduate students, not (trainee) psychotherapists.

To summarize, there is little research about ERA training targeting psychotherapists’ ERA. However, the two available studies show that psychotherapists’ ERA can be trained, at least the recognition for micro expressions and subtle emotional expressions and in the short-term. In studies II and III of this thesis, I evaluated two ERA trainings in a sample of undergraduate students (Study II) and a sample of trainee psychotherapists (Study III). The aim was to develop ERA training procedures that could be used in psychotherapy education.
Emotion recognition accuracy facets of interest

The present thesis focuses on two distinct but related ERA facets – multimodal ERA and micro expression ERA. In the following, I will explain those two facets and argue for why they are of special interest in the psychotherapy and psychotherapy education context.

Multimodal emotion recognition accuracy

*Multimodal ERA* describes the recognition accuracy for discrete non-verbal channels of expression/perception and the combination of those channels (e.g., facial, vocal and bodily cues). The existing studies about training non-verbal ERA in healthcare professionals include only facial expressions (often-times micro expressions) and they mostly use still pictures as stimuli. However, we know from previous research that emotions are communicated via different non-verbal channels, e.g., through bodily postures and tone of voice, beyond facial expressions and possible other channels (see, e.g., Bänziger, 2016). Assessing ERA in multiple modalities is relevant since people vary in their accuracy for different modalities of expression/recognition (see, e.g., Laukka et al., 2021; Schlegel, Boone, et al., 2017) and multimodal ERA tasks can provide a more differentiated picture of someone’s ERA. In the present studies, multimodal ERA was operationalized as the accurate perception of visual expressions (video-only), non-verbal auditory expressions (audio-only; solely based on prosody) and a combination of non-verbal audio-visual expressions (audio-video combined). Multimodal ERA can be assessed via standardized computerized tasks that include dynamic video and audio clips of actors or trained lay people displaying emotions. An advantage of using dynamic multimodal clips is that they present natural sequences of emotional expressions. The displays are, thus, more realistic and may be experienced as more ecologically valid than still pictures. Further, multimodal tasks allow for the assessment of other expression channels beyond mere facial expressions, which can help in establishing individual profiles for ERA and provides the opportunity to target single modalities, meaning that it is possible to individualize training based on each person’s individual needs. Assessing and training ERA for single modalities could also become useful for certain psychotherapeutic approaches and settings, e.g., specifically training auditory ERA for psychotherapists working with a lying-down setting in psychoanalysis, or for those providing telehealth interventions.

Micro expression emotion recognition accuracy

Very brief emotional expressions (in this case facial expressions) are considered micro expressions (see Ekman, 2003; Ekman & Friesen, 1969) when they last for 200ms or less (although some research suggests a longer duration; see
Most research about micro expressions understands them as involuntary cues of concealed emotions, i.e., remnants of emotional expressions that were, consciously or unconsciously, tried to squelch, neutralize or mask for various reasons (e.g., due to social desirability, lying, or other deception). The idea goes back to Darwin’s (1872) inhibition hypothesis suggesting that it is difficult to control facial muscles voluntarily, which might lead to involuntarily revealing true emotions (see also Ekman, 2009). Research indicates that it is hard to recognize micro expressions with an untrained eye (see e.g., Ekman & Friesen, 1969; Porter & ten Brinke, 2008; Frank & Svetieva, 2015), which is why it might be important to train micro expression ERA as part of psychotherapy education programs. When detected properly, micro expressions could provide important information about patients’ well-being and about risk assessment that psychotherapists could use for diagnosis and decision making (see, e.g., Ekman & Friesen, 1969; Frank & Svetieva, 2015). However, more recent research posits a more cautious understanding of micro expressions as indicators of (conscious) deception, like attempts to lie (Burgoon, 2018; Jordan et al., 2019; Porter and ten Brinke, 2008; Vrij et al., 2019; Weinberger, 2010; Zloteanu et al., 2021). Nonetheless, beyond deception, micro expressions could have the potential to provide useful information about conflicting, blended or confusing emotions that patients have, but don’t fully understand yet, or about repressed or dissociated emotions that could be explored in a safe therapeutic setting (see, e.g., Donovan et al., 2017).
Overview of the thesis

Aims and background

There is reason to believe that psychotherapists’ ERA is an important contributing factor to psychotherapy process and outcome, yet the literature about ERA assessed via standardized measures is still in its infancy (see Abargil & Tishby, 2021). Furthermore, research about psychotherapists’ ERA has been mixed. Some studies found that psychotherapists (Machado et al., 1999) and trainee psychotherapists (Pauza et al., 2010) were better at recognizing non-verbal emotional expressions than other populations, while other studies did not find differences between psychotherapists and matched controls (Hassenstab et al., 2007) or counseling trainees and undergraduate students (Hutchison and Gerstein, 2012). Not much is known about the impact of psychotherapy education and experience on ERA either (Machado et al., 1999). The studies in the field vary in their quality, but many made use of rather small samples and/or did not use control groups. Building on that, this thesis aims to fill some gaps in the literature by investigating trainee psychotherapists’ ERA and how this ability could be trained using standardized computerized training procedures as part of psychotherapy education.

In Study I, we investigated whether trainee psychotherapists were better at detecting non-verbal emotional expressions in others than a control group of undergraduate students, using two ERA tasks and a relatively large sample. Further, we followed those trainee psychotherapists from the beginning to the end of their psychotherapy education and assessed whether their ERA developed more positively due to the standard psychotherapy education, compared to the control group.

If psychotherapists’ ERA predicts psychotherapy outcome, it is relevant to implement interventions to improve ERA during psychotherapy education. This might be particularly relevant since previous research could not establish a clear answer to this question. In fact, the results of Study I suggest that standard (PDT and CBT) education in itself does not lead to significant increases in trainee psychotherapists’ ERA compared to a control group. Thus, my colleagues and I developed two standardized computerized trainings for ERA. One training targeted multimodal ERA and the other training targeted micro expression ERA. We evaluated the trainings in two randomized controlled studies (Studies II and III). In Study II, we evaluated the two ERA trainings in comparison to an active control group in a sample of undergraduate students.
to establish the general effectiveness of the trainings one week after the last of three (once-weekly) training sessions. In Study III, we administered the two trainings to a trainee psychotherapist population to test whether also trainee psychotherapists’ ERA could be improved using standardized computerized training. Further, in this study, we investigated whether the training effects were maintained one year later (follow-up).

Participants

For Studies I and III, trainee psychotherapists that attended Stockholm University’s clinical psychology program (psykologprogrammet) were investigated. Data was collected in terms 7 and 9 of their 10 terms long education. During those terms, the trainee psychotherapists train in either PDT or CBT (theoretically and practically) and conduct therapies under supervision at the university clinic. The interest in trainee psychotherapists – over fully educated, experienced psychotherapists – has several reasons. On the one hand, it was practical to recruit trainee psychotherapists at Stockholm University’s psychologist program because of the longitudinal design of the project. On the other hand, it also theoretically makes sense to assess trainees, as it is known from previous research that the empathy–outcome relationship is particularly strong for less experienced psychotherapists (Elliott et al., 2011), which could extend to ERA. Further, we included both PDT and CBT trainees, on the one hand to receive a decent sample size, and on the other hand, since we were interested in ERA as a therapist effect relevant in various forms of psychotherapy, in other words, as a possible common factor in psychotherapy (see Wampold, 2015). The sample sizes were not predefined, instead we aimed at including as many trainee psychotherapists as possible during three enrollment periods (three groups of trainee psychotherapists starting in term 7).

For Study II, an undergraduate student sample was included to first establish the general effectiveness of the ERA trainings.

Emotion recognition accuracy measures

The studies made use of the same ERA measures. The Emotion Recognition Assessment in Multiple Modalities test (ERAM; Laukka et al., 2021) was used as measure for multimodal ERA. It is a standardized computerized task that allows for assessing separate ERA scores for different modalities using dynamic clips of audio-only, video-only, and combined audio-video displays. In the audio modality, the participant is presented with an actor expressing an emotion in a pseudo-language ("ne kal i bam sud molen" and “kun se mina lod belam”) and, accordingly, needs to infer the emotional expression based on prosody alone. In the video modality, the participant needs to infer the
emotional expression based on facial expressions and upper body language alone, and in the audio-video modality, the participant is presented with a combination of those two modalities. The ERAM consists of 72 dynamic items taken from the Geneva Multimodal Emotion Portrayals Core Set (GEMEP; Bänziger et al., 2012) and assesses the twelve emotions anger, anxiety, despair, disgust, fear, interest, joy, pleasure, pride, relief, irritation and sadness. The items can be divided in positive and negative valence items, as well as by their arousal level. The items anger and irritation, for example, are varying in their level of arousal or intensity of expression. The participant’s task is to decide as quickly as possible which emotion was displayed with the help of a forced choice list of the twelve possible emotions. The ERAM was the primary outcome measure for the multimodal ERA training. Figure 1 shows a fictive picture (not taken from GEMEP) as an example for how a visual item of the ERAM could look like. The item was reenacted by an actor.

*Figure 1 Illustration of how a visual item of the ERAM could look like (fictive item, shown with permission by Lana Giese)*

Further, a self-developed standardized computerized micro expression recognition task (MICRO) that consists of 70 items was used to assess micro expression ERA. This task shows still pictures of actors’ faces that display seven basic emotions (happiness, surprise, fear, disgust, sadness, anger, contempt) for 200ms that are double-masked with neutral facial expressions for 2sec. This way, a micro expression is produced from still pictures of faces. The items are derived from the Radboud Faces Database (Langner et al.,
Also in this task, the participant has to indicate as quickly as possible which emotion was displayed using a forced-choice list of seven emotions. The MICRO was the primary outcome measure for the micro expression ERA training. *Figure 2* shows an example for how micro expressions are created via double-masked still pictures. The pictures are reenacted by an actor (not taken from the Radboud Faces Database).

*Figure 2 Illustration of the procedure to create micro expressions via still pictures (fictive items, shown with permission by Lana Giese)*

In *Studies II and III*, also a third ERA measure was included – the *Patient Emotion Cue Test* (PECT; Blanch-Hartigan, 2011). The PECT includes a combination of verbal and nonverbal displays of emotion using 47 video clips of a young female actor displaying the five emotions *anger, sadness, happiness, anxiety*, and *confusion*, as well as *neutral* expressions. The video clips (see *Figure 3* for a screenshot of a clip) are fictional scenes from medical interactions and include verbal statements, like “It’s just being gradually getting worse” or “Can I play golf again?”. The participant is to indicate as quickly as possible which emotion was displayed with the help of a forced-choice list of emotions. The PECT was used as an additional ERA outcome measure to explore whether the ERA trainings led to transfer effects for other ERA facets, in this case the recognition accuracy for a combination of verbal and nonverbal emotional expressions in medical settings, a setting related to psychotherapy.
Emotion recognition accuracy trainings

Study II and III used the same ERA trainings. There were two ERA trainings: one group conducted a multimodal ERA training and one group conducted a micro expression ERA training. Furthermore, we used an active control “training” for a third group. Each intervention was self-administered on a computer and took about 10-15 minutes to complete. The participants were instructed to train once weekly during a three-week period. In the first training session, the two ERA training groups also watched an informative video lecture (see Figure 4 for a screenshot) about emotional expressions that was about 15 minutes long. The active control training group did not watch the video.

There has been research about which components of ERA training might be most effective. Blanch-Hartigan (2012) found in an evaluation of a specific ERA training for medical contexts that comprehensive training is most effective and that training with feedback was the best single component for training. Simply watching videos of emotional expressions did not have a great impact. A meta-analysis (Blanch-Hartigan et al., 2012) on the effectiveness of trainings for perception accuracy (not exclusively ERA) suggests that the training approach is more important than the length of training. Practice with feedback has been found to be the most effective component of training (e.g., compared to practice alone or instruction alone) and comprehensive trainings were most effective. Following this, we adopted a training with feedback approach for the two ERA trainings, in addition to the information component (video lecture).
The multimodal ERA training made use of the same item database as the ERAM (Laukka et al., 2021), the GEMEP Core Set (Bänziger et al., 2012), but did not include items that were part of the ERAM (to avoid direct recollection effects). In the multimodal training, the first modality that is trained is the audio-video modality, as this is considered to be the easiest condition. As a progression, next, the video-modality is trained, and then the audio-modality. For each modality, the program randomly chooses 2 clips per emotion (24 emotional expressions per modality) out of a pool of 144 items. After each presentation, the participant is asked to indicate which emotion was expressed using a forced-choice list of the twelve emotions of the ERAM. In contrast to the ERAM, the participant receives immediate feedback on their performance (correct or incorrect) and what the correct would have been if they answered incorrectly. In the end of each training session, the participant even receives extended feedback about their typical errors in form of a confusion matrix (see Figure 5 for an example).
The micro expression ERA training follows the same principle as the MI-CRO in that it creates micro expressions from still pictures of faces via double-masking. However, the items are from another database, the Karolinska Directed Emotional Faces dataset (Lundqvist et al., 1998). Each session, the program chooses 60 items out of a pool of 336 items. Due to a coding error, anger was not trained and thus, we took this into account in the analyses in Studies II and III. After each presentation of a facial micro expression, the participant has to indicate which emotion was briefly shown and then receives immediate feedback on their performance and the correct answer in case of incorrect labeling. In the end of each session, they even receive extended feedback in form of their individual ERA rates per emotion (see Figure 6 for an example).

Figure 5 Example of extended feedback (multimodal ERA training)

Figure 6 Example of extended feedback (micro expression ERA training)
As active control training, an emotional working memory task with N-back format was used (see Gerhardsson et al., 2019). In this task, the participant is presented with positive, negative and neutral pictures from the International Affective Picture System (Lang et al., 2008) and has to indicate whether a certain picture had already been presented N-pictures before. See Figure 7 for an illustration of the principle behind the task. The original items may not be published and were, thus, replaced in the figure. This task also includes a phase with immediate feedback about performance in the beginning. This, and that the task was about emotionally laden pictures and, thus, had to do with emotions, led to the decision to use it as active control training. The task consists of four blocks á 72 items and takes about 15 minutes. There is no extended feedback about performance.

Figure 7 Example items of the control training (reproduced from Gerhardsson, 2022, p. 56, with permission)

Ethical considerations

All three studies were approved by the Swedish Ethical Review Authority (etikprövningsmyndigheten, dnr 2014/1265-31 and dnr 2015/1948-31) and all participants provided written informed consent for their participation. Studies II and III were preregistered at Open Science Framework (https://osf.io/3y2gb/).
The three studies had, in my view, little potential to distress or harm participants. On the contrary, we hoped that the ERA trainings could lead to improvements in the participants’ ERA, which could, theoretically, have positive influence on their interpersonal abilities and psychotherapist skills. Indirectly, improvements in ERA might even have positive effects on the trainees’ patients’ therapy results. Still, the assessment of ERA could also have led to dissatisfaction or self-criticism for some participants that perceived their own performance as low. For that reason, the test leader regularly checked in with the participants and support was offered after the posttest and the follow-up. Only some participants that dropped out of the study made use of this. The participants that dropped out volunteered that their discontinuation had to do with personal reasons and stress, which made it too burdensome for them to continue with the study, and that it had nothing with the topic of the study. This was, of course, respected by the researchers. Further, after the posttest in Study II and the follow-up in Study III, all participants were debriefed and the participants that were part of the control group were offered the opportunity to complete the real ERA trainings. However, none of the participants made use of this offer. All participants were also compensated for their participation with course credits or gift cards, and some even with sandwiches during lunch hours.

Research projects and funding

The three studies were financed by Stockholm University and research grants from the Marcus and Amalia Wallenberg Foundation (Marcus och Amalia Wallenbergs Minnesfond; grant no. MAW 2013.0130) and by Forte and the Marie Skłodowska-Curie Actions from the EU commission (Cofas FIIP-project, dnr 2013-02727).

The studies of this thesis were part of the abovementioned broader research projects about trainee psychotherapists’ ERA. The project funded by Forte and the Marie Skłodowska-Curie Actions project (primary investigator: Dr. Isabelle Letellier) investigated trainee psychotherapists’ ERA in the beginning and the end of psychotherapy education and how ERA might be associated with the outcome of the training psychotherapies that the trainees conduct during term 7 to 9 at the university outpatient clinic. Further, the project investigated how trainee psychotherapists experience the work with emotions in psychotherapy, using qualitative methods. The project funded by the Marcus and Amalia Wallenberg Foundation (primary investigator: Prof. Stephan Hau) investigated ERA trainings for trainee psychotherapists, and how ERA and ERA training might be associated with psychotherapy outcome and alliance. Furthermore, the project explored neural correlates of ERA and ERA training and the contribution of own facial expressivity to learning ERA; data to be presented in forthcoming articles.
Overview of the studies

**Study I:** Trainee psychotherapists are better at recognizing non-verbal expressions of emotions than a control group of undergraduate students, but they don’t improve in this competency significantly more than the control group after one and a half years of education.

**Study II:** Two ERA trainings significantly improve undergraduate students’ emotion recognition accuracy for multimodal expressions and facial micro expressions, respectively, in the one-week posttest compared to one another and to an active control training.

**Study III:** Trainee psychotherapists’ emotion recognition accuracy improves after explicit training in the one-week posttest, and unimodal emotion recognition accuracy even in the one-year follow-up, compared to an active control training.
Study I

Döllinger, L., Letellier, I., Högman, L. B., Laukka, P., Fischer, H., & Hau, S. (2023). Trainee psychotherapists’ emotion recognition accuracy during 1.5 years of psychotherapy education compared to a control group: No improvement after psychotherapy training. (*accepted by PeerJ, 14th of September 2023*)

Aims and background

Research about psychotherapists’ ERA has been mixed. Some studies found that psychotherapists (Machado et al., 1999) and trainee psychotherapists (Pauza et al., 2010) were better at recognizing non-verbal emotional expressions than other populations, while other studies did not find differences between psychotherapists and matched controls (Hassenstab et al., 2007) or counseling trainees and undergraduate students (Hutchison and Gerstein, 2012). Machado et al. (1999) found that practicing psychotherapists were better at recognizing non-verbal displays of emotion compared to undergraduate psychology students who intended to train as psychotherapists in the future, but apart from this study, it is unclear whether standard psychotherapy training can lead to better ERA. Based on that, we had two research questions in Study I. First, we investigated whether trainee psychotherapists had higher ERA than a sample of undergraduate students at baseline. Our hypothesis was that the trainee psychotherapists would be superior in ERA, possibly due to heightened interest in socio-emotional variables in interpersonal exchanges or other traits that could be associated with their choice of profession. Then, we investigated whether the trainee psychotherapists’ ERA increased more strongly than the control groups’ after one and a half years of psychotherapy training due to the theoretical and practical education and the work with patients at the university clinic. Our hypothesis was that the trainee psychotherapists would display larger gains in ERA than the control group.
Method

In the beginning of the study, 49 trainee psychotherapists ($M_{age} = 30.35$, $SD_{age} = 7.15$; 34 PDT, 15 CBT) and 105 undergraduate students ($M_{age} = 25.82$, $SD_{age} = 7.78$) enrolled in the study and completed the ERAM and MICRO ($N = 154$). The age difference between the groups was significant ($U = 3774$, $p < .001$, 95% CI [3.00, 6.00]; $r = .38$. 95% CI [.25, .51]), but did not influence ERA. The trainee psychotherapists were in the beginning of their psychotherapy education (term 7) but had already attended six terms of psychology education. For the first research question about the difference in ERA between trainee psychotherapists and the undergraduate students, we performed independent $t$-tests and Wilcoxon-Mann-Whitney $U$-tests.

After one and a half years, when the trainee psychotherapists were in the end of their psychotherapy education (term 9), all participants were invited back to the laboratory. At follow-up, 31 trainee psychotherapists ($M_{age} = 31.77$, $SD_{age} = 7.62$; 20 PDT, 11 CBT) and 41 undergraduate students ($M_{age} = 28.00$, $SD_{age} = 9.79$) came back to complete the ERAM and MICRO once again ($n = 72$). The age difference between the groups was still significant ($U = 872$, $p < .01$, 95% CI [1.00, 8.00]; $r = .38$. 95% CI [.25, .51]), but did not influence ERA or the ERA change scores. We also explored reasons for dropout, but only found that the likelihood of dropout increased slightly the younger the participants were. For the second research question about the differences in between-group ERA change trajectories, we performed mixed multilevel modeling (see, e.g., Field & Wright, 2011; Finch et al., 2019) with maximum likelihood estimation for missing data (Enders, 2022), and then once again without missing data handling method to verify our findings.

Results

We found that the trainee psychotherapists in the beginning of psychotherapy education were significantly better at recognizing multimodal ($t(152) = 3.64$, $p < .001$, 95% CI [0.25, 0.86]) and micro expressions ($t(152) = 2.02$, $p = .05$, 95% CI [0.00, 0.10]) than the control group. The standardized effect sizes were moderate ($d = .63$, 95% CI $[0.32, 0.98]$) for the ERAM and small ($d = .35$, 95% CI $[0.03, 0.76]$) for the MICRO. Figure 8 contains violin plots displaying the distribution of the data, means and confidence intervals. The ERAM 95% confidence intervals of the groups do not overlap. However, they overlap slightly for the MICRO. When exploring the ERAM modalities separately, we found that the advantage of the trainee psychotherapists can be seen in the audio and the audio-video, but not in the video modality. Further, the trainee psychotherapists were superior even for positive and negative valanced items and for high and low arousal items.
The mixed multilevel analyses revealed that the ERA change trajectories of the two groups differed significantly, for multimodal ERA (between-group difference in slope = -.04, SE = .02, t(70) = -2.37, p = .02; 95% CI [-0.08, -0.01]) and for micro expression ERA (between-group difference in slope = -.09, SE = .03, t(70) = -2.54, p = .01; 95% CI [-0.16, -0.02]). However, the effects were not in line with the hypothesis. See Figure 9 for a visual display of the ERA changes. The trainee psychotherapists did not improve in multimodal ERA from pretest to follow-up (t(70) = -0.18, p = .85, 95% CI [-.03, .03]), while the control group’s improvement of 5% points was significant (t(70) = 3.88, p < .001, 95% CI [-.12, .01]). The difference of the between-groups slopes was moderately sized (d = .63). When it comes to micro expression ERA, both groups improved significantly from pretest to follow-up, but the control group’s improvement was steeper. The trainee psychotherapists improved by 6% points (t(70) = -2.43, p = .02, 95% CI [-.12, -.01]), whereas the control group by 15% points (t(70) = -6.87, p < .001, 95% CI [-.20, -.11]). The difference of the between-groups slopes was moderately sized (d = .57). The sensitivity analyses after excluding all dropouts confirm that the trainee psychotherapists did not profit from psychotherapy education when it comes to improved ERA. The differences between the groups were even less pronounced. Generally, there were no differences between the PDT and CBT students’ ERA.

Figure 8 Study I: Violin plots of ERA per group at pretest

Note: The violin plots display the kernel probability density of the data at pretest with means (points) and 95% Confidence Intervals.
Discussion and conclusion

Study I shows that trainee psychotherapists in the beginning of psychotherapy education are better at recognizing non-verbal emotional expressions in others than a control group of undergraduate students. This was true for multimodal expressions and micro expressions. However, there is indication to believe that this effect is particularly strong for multimodal ERA, and especially in the modalities that involve auditory ERA. The fact that the trainees were better in all valence and arousal categories further strengthens the results for the ERAM. At this point in time, the trainee psychotherapists were in the beginning of their psychotherapy education. Thus, we don’t think that the higher multimodal and micro expression ERA of this group had to do with clinical experience or education. Instead, the results cautiously suggest that the trainee...
psychotherapists might be better at recognizing non-verbal emotional expressions on a trait level. To be able to say this with more certainty though, there need to be further longitudinal studies investigating the trainee psychotherapists from the beginning of their education (term 1) or even before that point, as we cannot preclude that the psychology education from term 1 to 6 might already have led to ERA improvements. We suggest that the trainee psychotherapists’ higher ERA might even be involved in their choice of profession, but this hypothesis needs to be tested in future studies as well.

However, after one and a half years of psychotherapy education, the trainee psychotherapists had not improved in multimodal ERA, while the control group did; and not significantly more than the control group in micro expression ERA. This suggests that standard PDT and CBT psychotherapy education does not necessarily lead to higher ERA. The main implication of this finding, in my opinion, is that socio-emotional competencies, like the recognition accuracy for non-verbal emotional expressions, need to be explicitly addressed or trained in the psychotherapy education to substantially improve. This finding lays the basis for Study III in which we investigated whether ERA training would lead to immediate and long-term improvements for trainee psychotherapists. Another possible reason for this finding could be that the ERA facets that are relevant in psychotherapy, and that are addressed in psychotherapy education and practice, might not be the same as the facets assessed with standardized computerized tasks as the ERAM and MICRO. The main limitations of Study I are the high dropout rate from pretest to follow-up (second research question), the relatively low Cronbach’s alpha values of the ERAM and the fact that we cannot preclude that earlier improvements of ERA due to general psychology education could have taken place before the start of the study.
Study II


Aims and background

In *Study II*, we evaluated two standardized computerized ERA trainings that my colleagues and I developed. For that purpose, we first used a sample of undergraduate students. One training was for multimodal ERA and the other training was for micro expression ERA. It is unclear whether ERA improves simply as a result of experience and psychotherapy education. Machado et al. (1999) found that experienced psychotherapists had better ERA for non-verbal displays of emotions than inexperienced trainees, which might point towards ERA improving with experience. On the other hand, using a mixed research design, *Study I* did not find improved ERA after psychotherapy education in comparison to a control group. Thus, the goal was to create ERA trainings that could be used as part of psychotherapy education and to explore whether they are effective. However, first, the two trainings were tested in a broader group of undergraduate students to establish their general effectiveness.

Method

The participants in *Study II* were 72 undergraduate students (*M*<sub>age</sub> = 24.69, *SD*<sub>age</sub> = 7.69, range = 18–51 years, 54 women), of which 67 also participated in the three training sessions and in the posttest (*M*<sub>age</sub> = 24.45, *SD*<sub>age</sub> = 7.38, range = 18–51 years, 49 women). The participants were randomly assigned to either the multimodal ERA training, the micro expression ERA training or the active control training (stratified for gender). There were no age differences between the groups. In the pretest and the posttest, the participants completed the three ERA tasks (ERAM, MICRO, PECT). Between pretest and posttest, they trained once a week during a three-week time period. The trainings took place at the laboratory of the university and the participants were assisted by a research assistant. Whereas the participants were blind to their condition, the
research assistants were not (single-blind study). The posttest took place one week after the last training session.

Since the conditions for parametric statistical analyses were not fulfilled, we analyzed the data with a, for psychological research, rather unconventional non-parametric analysis called Aligned Rank Transform (Wobbrock et al., 2011) for mixed factorial designs. This analysis is a modified version of the Rank Transform that allows to test for interaction effects. For this analysis, the data is aligned, stripped of the effect of other variables (main effects or interaction effects, based on which effect is estimated), and then rank transformed. We applied a 2x3 mixed factorial design and allowed for a random intercept.

**Results**

The Aligned Rank Transform mixed ANOVAS showed significant *time by training* interactions for the ERAM \((F(2,64) = 6.83, p < .002)\) and the MICRO \((F(2,64) = 10.18, p < .001)\). Post-hoc contrast analyses revealed that the pretest–posttest difference of the multimodal training group \((\text{diff} = .15)\) was significantly larger than the pre–post differences of the other two groups, which were both improvements of 6% points (compared with micro expression training: \(\chi^2 (1, N = 44) = 9.06, p = .005\); compared with control training: \(\chi^2 (1, N = 44) = 11.57, p = .002\)). Further, the pretest–posttest difference of the micro expression training group \((\text{diff} = .26)\) was significantly larger than the pre–post differences of the multimodal training group \((\text{diff} = .06, \chi^2 (1, N = 44) = 19.50, p < .001)\) and the control training group \((\text{diff} = .12, \chi^2 (1, N = 46) = 8.78, p = .01)\). The effects can also be observed by the non-overlapping, respectively overlapping, interquartile ranges (Figure 10). The within-group improvements in the training groups were much larger than for the other two groups for each respective outcome measure. For the ERAM, the effect of the multimodal training was very large \((d_z = 2.04)\), and moderate for the micro expression training \((d_z = .64)\) and the control training \((d_z = .71)\). For the MICRO, the effect of the micro expression training was very large \((d_z = 1.33)\), moderate for the multimodal training \((d_z = .54)\) and small for the control training \((d_z = .40)\).

We also found a significant interaction effect when considering only the audio modality of the ERAM \((F(2,64) = 4.47, p < .02)\). The pretest–posttest difference for the multimodal training group \((\text{diff} = .16, d_z = 1.39)\) was significantly larger than for the other two groups (compared with the micro expression training: \(\text{diff} = .06, d_z = .41, \chi^2 (1, N = 44) = 5.23, p = .04\); compared with the control training: \(\text{diff} = .04, d_z = .30, \chi^2 (1, N = 44) = 8.05, p = .01\)). There were no interaction effects for the video-only or the audio-video modality of the ERAM test. Further, there was no interaction effect for the third ERA outcome, the PECT. The correlational patterns between ERA baseline scores and
pre-post change scores suggest that low baseline could be associated with higher improvements.

Discussion and conclusion

In Study II, we could show that the standardized computerized ERA trainings effectively improved the participants’ ERA from pretest to posttest one week after the last training session. The within-group changes for the multimodal ERA training and the micro expression ERA training were both large. After establishing the effectiveness of the trainings, they could now be applied and evaluated in other populations, like trainee psychotherapists or other professions that rely on socio-emotional perception skills, like emotion recognition. Interestingly, the effect for the ERAM seemed to be driven by the audio-only modality. Particularly the recognition of unimodal audio stimuli improved in response to the multimodal training. Since the audio modality had the lowest accuracy rates, it is very encouraging that this ability can be successfully trained. In fact, there is reason to believe that low baseline ERA in general was associated with larger improvements. This is encouraging, as it suggests that training could be especially helpful for low performers or for more difficult modalities. However, there might simply be more room for improvement in those instances. This needs to be further explored in future studies. Furthermore, the fact that there were no transfer effects for verbal and nonverbal ERA in medical settings suggests that ERA training needs to be specific to be effective.
Figure 10 Study II: Violin plots with box plots of the training effects for the three ERA measures

Note. The violin plots display the kernel probability density of the data at the different values for the three training groups. The box plots within the violin plots include the median (-) and the interquartile range (box), as well as the minimum and maximum (whiskers). Means were added in form of rhombuses. Small black dots display outliers.
Study III


Aims and background

*Study III* aimed to replicate the findings of *Study II* in a sample of trainee psychotherapists, as this was the target population of this thesis. *Study II* could show that the multimodal and the micro expression ERA training led to significant improvements in ERA one week after the last training session. In *Study III*, it was also investigated whether the effects were long-lasting (one-year follow-up).

Method

Sixty-eight trainee psychotherapists ($M_{\text{age}} = 30.9, SD_{\text{age}} = 6.18$; 39 PDT and 29 CBT) participated in the pretest (ERAM, MICRO, PECT) in the beginning of their psychotherapy education (term 7) before being randomly assigned (stratified for gender and therapy approach) to one of the two ERA trainings or the active control training. The participants trained on a computer stationed at the university clinic after having received information about how to conduct their respective training. They trained independently, but received regular reminders and offers of help by the main test leader. The test leaders that assisted with the tests at the laboratory were blind to the participants’ conditions. Sixty trainee psychotherapists ($M_{\text{age}} = 30.1, SD_{\text{age}} = 6.23$; 36 PDT and 24 CBT) participated in the posttest one week after the last training session and 55 ($M_{\text{age}} = 29.84, SD_{\text{age}} = 6.21$; 33 PDT and 22 CBT) also in the follow-up measurement one year later. To investigate whether the multimodal ERA training was leading to steeper improvements in multimodal ERA and the micro expression ERA training to steeper improvements in micro expression ERA, we performed mixed multilevel modeling analyses with maximum likelihood estimation. We investigated the short-term training effect as the ERA change from
pretest to posttest one week after the last training session. The maintenance
effect was measured by the ERA change from pretest to follow-up (which took
place one year after the posttest). We also explored whether there were any
transfer effects of the trainings for verbal and non-verbal ERA in medical set-
tings (PECT) or any influences of affective state and ERA baseline on ERA.

Results

The multimodal ERA training showed larger pretest–posttest improvements
in multimodal ERA than the micro expression training (between-group differ-
ence in slope = -0.09, SE = .03, t(109) = -3.36, p < .001; 95% CI [-0.15, -0.04])
and the control training (between-group difference in slope = -0.13, SE = .03,
t(109) = -4.69, p < .001; 95% CI [-0.18, -0.08]). The differences in change
trajectories from pretest to posttest were large (multimodal vs. micro expres-
sion training: \( d = .90 \)) to very large (multimodal vs. control training: \( d = 1.63 \)).
The within-group change in the multimodal training group was 15% points,
whereas the within-group change in the micro expression training group was
6% points. The control training group did not improve from pretest to posttest.
Between pretest and follow-up, there were no significant between-groups dif-
fferences in slopes. See Figure 11 for a visual display.

Figure 11 Study III: ERAM change for the three training groups

Note. N = 68. Based on estimated marginal means. Error bars represent 95% Confidence Intervals.
Since the ERAM allows for the assessment of unimodal ERA, we performed the analyses also for the separate modalities. For the audio-only and the video-only modality, there were significant large to very large differences in slopes from pretest to posttest in comparison to both the micro expression training group (between-group difference in slope_{audio} = -.11, SE = .04, t(109) = -2.80, p = .01; 95% CI [-0.19, -0.04], d = 1.00; between-group difference in slope_{video} = -.10, SE = .04, t(109) = -2.55, p = .01; 95% CI [-0.18, -0.02], d = .83) and to the control training group (between-group difference in slope_{audio} = -.17, SE = .04, t(109) = -4.27, p < .001; 95% CI [-0.25, -0.09], d = 1.80; between-group difference in slope_{video} = -.13, SE = .04, t(109) = -3.25, p < .001; 95% CI [-0.21, -0.05], d = 1.3). In addition, in the unimodal channels, even the pretest–follow-up differences in slopes were significant and of large size, but only when comparing the multimodal training to the control training, not to the other ERA training (between-group difference in slope_{audio} = -.11, SE = .04, t(109) = -2.54, p = .01; 95% CI [-0.19, -0.03], d = 1.00; (between-group difference in slope_{video} = -.09, SE = .04, t(109) = -2.04, p = .04; 95% CI [-0.17, -0.00], d = .90). However, in the combined audio-video modality, neither the pretest–posttest, nor the pretest–follow-up between-groups differences in slopes were significant. See Figure 12 for a visual display.
Figure 12 Study III: ERAM change for the three training groups per modality

The micro expression training led to a significantly steeper micro expression ERA improvement from pretest to posttest than the active control training (between-group difference in slope = -0.11, SE = .06, \( t(109) = -1.96, p = .05; 95\% \text{ CI}\ [-0.22, -0.00], d = .69 \)). All within-changes were significant. The micro expression training group improved with 24% points, the multimodal training group with 17% points and the control training group with 13% points. There was no detectable difference in slopes compared to the multimodal training. From pretest to follow-up, there were no significant differences in ERA slopes between the groups. See Figure 13 for a visual display of the results.

Note. Based on estimated marginal means. Error bars represent 95\% Confidence Intervals.
The exploratory analyses regarding the third, independent outcome measure for verbal and non-verbal ERA in medical settings (PECT) did not show any interaction effects. There was no difference in change trajectory between either training group or the control group for the PECT, suggesting that there were no transfer effects for this specific ERA facet. Additional analyses showed that ERA baseline generally influenced the within-groups ERA change scores. For multimodal ERA, low baseline scores even predicted whether someone would respond to the training more strongly (higher pretest–posttest change). Affective state seemed to influence all groups’ ERA equally.

Discussion and conclusion
In this randomized controlled double-blind study (*Study III*), we could show that the multimodal ERA training and the micro expression ERA training were effective in improving trainee psychotherapists’ ERA from pretest to posttest one week after the last training session, compared to an active control training. The effect size for the micro expression training was moderate. The multimodal training on the other hand showed large to very large effects. Further, the multimodal training even proved to be superior to the other ERA training
from pretest to posttest (not only in comparison to the group that did not train in ERA). This suggests that trainee psychotherapists can successfully improve their micro expression ERA, and especially their multimodal ERA – at least in the short term – using standardized computerized training procedures. There were no lasting effects of the trainings from pretest to follow-up after 1.5 years when investigating general multimodal and micro expression ERA.

However, examining the different modalities of the ERAM suggests that the effects of the multimodal training were driven by the two unimodal conditions (audio-only and video-only), as the superiority of the multimodal training group towards the other two groups from pretest–posttest could only be observed in the unimodal presentations, but not in the combined modality (audio-video). Furthermore, from pretest to follow-up, the participants that trained in multimodal ERA still showed superior improvements in unimodal auditory ERA and unimodal visual ERA compared to the control training group, but, again, not in the combined modality. This suggests that, despite there being no long-term training effects for multimodal ERA in general, the multimodal training seems to lead to long-lasting improvements in ERA for unimodal expressions of emotion. To my knowledge, this is the first study showing that multimodal ERA training has long-lasting effects. That these effects only occur for the unimodal channels is likely due to the fact that the ERA for those modalities is more difficult and might thus respond more positively to training.

Further, low baseline ERA led to higher multimodal ERA improvements, suggesting that trainee psychotherapists with low baseline multimodal ERA could profit from the training most. This could also be a ceiling effect, meaning that the trainee psychotherapists with high baseline did not have much room for improvement. But still, multimodal ERA training might be beneficial especially for those trainees who struggle with multimodal ERA to begin with. The fact that there were no transfer effects of the ERA trainings to the measure about verbal and non-verbal emotional expressions in medical settings indicates that ERA training needs to be very specific to lead to significant and hopefully lasting effects.
Discussion

Summary of results

This thesis investigated trainee psychotherapists’ emotion recognition accuracy (ERA) and how it changes from the beginning to the end of 1.5 years psychotherapy education compared to a sample of undergraduate students. Further, it investigated the effectiveness of two ERA trainings, both in an undergraduate student sample and in a trainee psychotherapist sample.

In Study I, it was found that the trainee psychotherapists showed moderately greater multimodal ERA and slightly greater micro expression ERA than the undergraduate students in the beginning of psychotherapy education. However, the trainee psychotherapists’ multimodal ERA did not improve from the beginning to the end of the psychotherapy education, whereas the control groups’ improved with 5% points and the difference in slopes was significant and of moderate size. Further, there was a significant increase in micro expression ERA for the trainee psychotherapist group (6%), but the increase of the control group was steeper (15%) and the difference between those improvements was significant and of moderate size. This suggests that trainee psychotherapists’ ERA does not improve in any meaningful fashion in response to attending the 1.5 years psychotherapy education (at Stockholm University). There were no differences between the PDT and CBT students’ ERA.

Based on the finding that ERA does not improve simply by learning and practicing psychotherapy, we suggest that explicit training of ERA might be useful for psychotherapy education. For that reason, my colleagues and I developed two training programs for improving multimodal ERA and micro expression ERA. The trainings are computer-based and consist of dynamic (multimodal training) or still picture (micro expression training) stimuli of emotional expressions and the task to judge which emotion was expressed. Then the participants receive immediate and extended feedback about their performance. The trainings consist of three sessions (one per week) and can be administered autonomously. For investigating the effectiveness of the trainings, we compared the two training groups with each other and with a control group that received an active control training. The trainings were first evaluated in a sample of undergraduate students (Study II). The results show that the two trainings were effective in improving the participants’ ERA in the one-week
posttest. The participants that received the multimodal training became significantly better than the other two groups with an average improvement of 15% points and a very large standardized effect size, whereas the micro expression training and the control training both only improved with 6% points. A similar pattern could be shown for the micro expression training. The participants in this group showed significantly stronger improvements than the other two groups. Their within-person change was 26% points, whereas the other groups improved with 6% points (multimodal training group) and 12% points (control training group). The within-change can be interpreted as very large both according to the standardized effect size and relative to the other groups. Neither of the trainings showed any effect on the ability to recognize verbal and nonverbal emotional expressions in the medical clinical context (PECT outcome measure).

After establishing that the trainings were effective, we evaluated the same trainings in a group of trainee psychotherapists (Study III). Data was collected in the beginning of their psychotherapy education (term 7) and then again one week after the last training session. In this group, we also included a long-term follow-up about one year after the posttest (term 9). The results of the previous study were replicated and extended. There were large to very large differences in change trajectories for multimodal ERA from pretest to posttest in favor of the multimodal training. The multimodal training group improved with 15% points, the micro expression group with 6% points and the control training group not at all. Further, for micro expression ERA, the micro expression training group’s change trajectory of 24% points was significantly and moderately steeper than that of the control group (13% points), but there was no statistical difference to the multimodal training group that improved with 17% points. All within-group changes were significant. In the one-year follow-up, we first could not detect any long-term effects of either training. However, when considering the ERAM modalities separately, there is indication to believe that the multimodal training group was still superior in terms of being able to correctly recognize unimodal (audio-only and video-only) expressions one year later in comparison to the control group (but not compared to the micro expression training group). Like in the previous study, neither of the trainings affected verbal and nonverbal ERA in the medical clinical context (PECT outcome measure).

Similarities and differences of the results

When comparing the two studies that included trainee psychotherapists, it can be noticed that the results for the control group of Study III (those that did not receive an ERA training) generally support the result of Study I that the trainee psychotherapist’s ERA did not meaningfully improve in the course of the psychotherapy education. Both studies found very similar baseline values of 45% (Study I) and 46% (Study III) recognition accuracy for multimodal
expressions and 50% (Study I) and 53% (Study III) recognition accuracy for micro expressions. Further, Study III found no significant improvements in the one-year follow-up, suggesting that those participants that did not train in ERA did not improve in this capacity during the psychotherapy education. However, it must be noted that the follow-up in that study happened some months earlier than in Study I. Also, it needs to be noted that there still was a significant improvement in micro expression ERA in Study I. However, the confidence intervals overlapped and the control group of undergraduate students in this study improved as well (and even more strongly), suggesting that the improvement could not be attributed to the psychotherapy education.

Interestingly, the improvements due to the trainings were very similar for the undergraduate student (Study II) and the trainee psychotherapist samples (Study III). The multimodal training groups improved with 15% points in multimodal ERA in both samples, while the micro expression training group improved with 6% points in both samples. The control group only improved in the undergraduate sample, not in the trainee psychotherapist sample. Similarly, the micro expression training led to improvements of 26% points (Study II) and 24% points (Study III) in micro expression ERA, respectively, while the control group improved with 12% and 13% points, respectively. However, there was a different pattern for the multimodal training group that improved in micro expression ERA with 6% points in Study II, but with 17% points in Study III, even if they did not train micro expression recognition. This could cautiously be interpreted as a small transfer or spill-over effect of the multimodal training, but only in the trainee psychotherapist sample, and, thus, should not be overstated. The similarities in the patterns of improvement due to the two ERA trainings suggest a robustness of the trainings’ effects and that they can be applied to several populations.

In both training studies, there was indication to believe that the audio modality plays a special role in explaining the effectiveness of the multimodal ERA training. In Study II, interaction effects (pretest–posttest) were found for the audio modality and the ERAM total score, meaning that the participants that trained in multimodal ERA improved in the audio-only modality and the total score of the three modalities (but not in the video-only and the audio-video combined modality). In Study III, there were significant interaction effects for the ERAM total score and the two unimodal channels (audio-only and video-only) in the pretest–posttest comparisons, but not for the combined modality. In the pretest–follow-up comparison, the long-term effects of the multimodal training could only be seen for the unimodal channels. These findings suggest that the training effects are especially visible for auditory ERA (even if not only auditory ERA) and that the participants in the multimodal training group particularly profit from the training when it comes to auditory ERA or unimodal ERA in general. The effectiveness of the multimodal training may be driven largely by the audio modality (even if not exclusively).
These were, to my knowledge, the first studies investigating unimodal, auditory ERA in ERA training contexts.

Both training studies also explored whether baseline ERA might be associated with steeper improvements due to training. Previous studies speculated that there could be such a link since low baseline ERA might leave more room for improvement (see Blanch-Hartigan et al., 2012; Schlegel et al., 2017). *Study II* found negative correlations between ERA baseline and general pretest–posttest change scores, and *Study III* found that low ERA baseline predicted larger general ERA changes. Further, *Study III* also suggests that low multimodal ERA at baseline predicts that a participant responds to the training more strongly, but the sample size for this exploratory analysis was small and the result was not found for micro expression ERA. Variables like age, gender or psychotherapy approach did not influence ERA and ERA changes in the present studies.

**Discussion of the results and implications**

This thesis suggests that trainee psychotherapists’ multimodal and micro expression ERA is higher than undergraduate students’, but that it does not improve due to standard psychotherapy education. Further it suggests that multimodal and micro expression ERA can be successfully trained using computerized standardized training programs in the short term and the recognition accuracy for unimodal (audio or video) dynamic expressions likely also in the long term. These findings and their implications need to be discussed.

**Trainee psychotherapists’ ERA is higher than undergraduate students’**

Previous research did not provide a clear answer to the question whether (trainee) psychotherapists were better at recognizing emotional expressions in others compared to control groups that are not working psychotherapeutically. This thesis suggests that trainee psychotherapists indeed are better at both multimodal and micro expression recognition than undergraduate students, and, thus, supports the findings of Pauza et al. (2010) that compared psychotherapy trainees using different approaches with trainees of other professions. The present thesis and Pauza et al.’s study had larger samples than Hutchison & Gerstein (2012) and Hassenstab et al. (2007), who did not find differences between trainee psychotherapists and undergraduate students, respectively between practicing psychotherapists and matched controls. This might have made it easier to detect significant effects. Using a naturalistic measure for ERA, Machado et al. (1999) found that experienced psychotherapists were better at detecting non-verbal displays of emotion than undergraduate clinical...
psychology students. This suggests that experienced psychotherapists are better at recognizing non-verbal emotional expressions in others and that therapy experience and/or education might contribute to larger ERA. However, this was a cross-sectional finding. It is worth noting that Machado et al. found a similar result as the present thesis when it comes to the group difference between psychotherapists and undergraduate students, even though Machado et al.’s study investigated already fully educated therapists. This is positive, as the two studies apply very different methodologies (standardized ERA task and a measure based on video-recorded therapy session, respectively). However, it also contradicts the finding of the present thesis that psychotherapy education – or clinical experience, this question still remains open – leads to higher ERA. Again, this could have been due to methodological differences, being that Machado et al. used a naturalistic ERA measure.

Given that the trainee psychotherapists in the present thesis were in the beginning of their psychotherapy education, I suspect that the differences in ERA compared to the undergraduate non-clinical psychology students might have something to do with traits or abilities that trainee psychotherapists possess, although more research is needed to identify those factors. Their superior ERA could be influenced by traits that are associated with choosing to become psychotherapists. Maybe this population shows special interest in other people’s moods and emotional expressions in general, which might have influenced their choice of education. Also, motivational factors that might be closely linked to the profession or choice of profession might be involved, e.g., the intrinsic need for or interest in understanding others’ emotions. Maybe the trainee psychotherapists had a “headstart” compared to the undergraduate students, as they were already very interested in other people’s emotions. Given the design of the study, it also cannot be precluded that other educational or intellectual factors might have influenced the result as well. Maybe the trainee psychotherapists already had learned more about interpersonal dynamics, emotions or psychotherapy in general during their studies than the control group did, previous to starting to learn and practice their chosen therapy approach. Since the trainee psychotherapists were older and already had attended six terms of psychology classes, it is very well possible that their ERA might already have been positively affected by the education, in comparison to the control group that was on average younger and mostly read introductory classes in psychology. The length of education, other learning effects or specific personality traits could have played a role as well. Even though there were significant age differences between the groups, we found that age did not predict ERA or ERA changes. Still, developmental factors related to age could possibly have played in as well. Future research is needed to understand the group differences in ERA in more depth.
No changes in ERA from beginning to end of psychotherapy education

Although the trainee psychotherapists were very good at detecting non-verbal emotional expressions already in the beginning of their theoretical and practical psychotherapy education, this thesis suggests that their ERA does not improve due to standard PDT or CBT education, at least not due to the kind of education that we provide at Stockholm University. On the one hand, one could say that this is not very surprising, as neither approach specifically included any form of explicit ERA training, e.g., in the form of standardized computerized training or deliberate practice for ERA. So, explicit training of ERA is probably needed to meaningfully improve in this ability. This might be particularly the case for facets of ERA that are very specific, like micro expression recognition or the recognition of unimodal (especially auditory) expressions of emotion. These expressions might not even be consciously noticed (in the case of micro expressions) or are seldomly trained implicitly (e.g., in the case of prosody). Furthermore, maybe psychotherapy education focuses too much on verbal content – on what is communicated – above how emotional communication takes place non-verbally via body-language, facial expressions and prosody. Even though both approaches talk about patients’ emotions in supervision and to some extent also in the theoretical education, it might not be enough to stimulate greater changes. This might be different for other, more emotion-focused psychotherapy approaches. It would be worth investigating whether, e.g., Intensive Short-Term Psychodynamic Psychotherapy (Davanloo, 2001), Accelerated Experiential-Dynamic Psychotherapy (Fosha & Yeung, 2006), Emotion-focused Therapy (Greenberg, 2015), or other approaches that explicitly work with patients’ emotional shifts and the somatic experiencing of emotions in session, might lead to different results. Furthermore, approaches that make use of deliberate practice for working with patients’ emotions, like in Emotion Focused Therapy (see Goldman et al., 2021), might have greater success in terms of ERA improvements.

On the other hand, maybe standardized ERA tasks, like the ERAM and the MICRO, are not capable of capturing improvements in emotion recognition competencies that implicitly improve during the course of psychotherapy education. There might be a lot of socio-emotional learning taking place during those 1.5 years, but it might be limited to perceptual competencies related to psychotherapy situations and might not be transferred to standardized ERA tasks. Meaning, maybe the ERAM and MICRO did not measure changes in ERA that are relevant for psychotherapy. This puts into question the validity of standardized computerized ERA measures for this purpose, as well as the external validity of the results of this thesis. Maybe this is why Machado et al. (1999) found that clinical experience was associated with higher ERA using a naturalistic ERA measure, whereas the present thesis did not. Further, the psy-
chotherapy education includes a multitude of theories, methods and techniques and (hopefully) stimulates a vast variety of socio-emotional, reflective and interpersonal competencies. The attempt to investigate ERA as an isolated phenomenon or capacity might be challenging and too limited. For example, it is very possible that ERA for verbal displays and especially ambiguous verbal displays might improve due to education (even though this needs to be tested in research). Also, it is possible that it is less important that the psychotherapist is able to accurately read a patient’s expressed emotion than that they can indicate that some emotional shift is happening in the patient, in other words, maybe the psychotherapist does not actually need to have perfectly accurate emotion recognition competencies, but needs to be a safe space and supportive to let the patient explore their emotions themselves. Just like Rogers (1975) redefined his conception of empathy to be a process of empathically meeting another person and “being sensitive, moment-to-moment, to the changing meanings which flow in this other person” (p. 4), describing empathy less as an accuracy, but more as an attitude to possess, something similar might be applicable to ERA. Topics like facilitating dyadic synchrony (Atzil-Slonim et al., 2023) might also be non-verbal therapeutically relevant skills that are not assessed via measuring ERA.

The results that trainee psychotherapists were relatively good at ERA and that psychotherapy education did not lead to meaningful improvements in trainee psychotherapists’ ERA has two mayor implications. The first one might be about how we select individuals for the psychotherapy education. I would suggest that individuals that attend the psychotherapy education on average are quite good at recognizing emotional expression in others, even though we don’t, yet, know why this is the case. But also in this population, there are individual differences in ERA. In Sweden, as well as in many other countries, trainees for the clinical psychology program are chosen largely based on academic merits. However, this might be a too limited indicator for suitability for this program. Since ERA does not seem to meaningfully improve due to standard education, I suggest that measures of individuals’ ERA could be used to select the most suitable candidates for the psychotherapy education. Obviously, I do not suggest that ERA should be the only indicator for acceptance into the program, but it could be one of many, alongside academic and intellectual merits and other socio-emotional, reflective and interpersonal variables linked to positive psychotherapy results and relationship building, like the capacity to establish and repair alliances and empathy (Heinonen & Nissen-Lie, 2020; Wampold, 2015) or possibly mentalizing capacities. Considering how we select suitable candidates for this profession becomes even more relevant in the light of studies about experience and therapy outcome. Longitudinal studies about psychotherapist experience and its relationship to outcome have produced inconclusive results. Goldberg, Rousmaniere, et al. (2016) found that psychotherapists achieved slightly worse outcomes with increasing experience; Goldberg, Babins-Wagner, et al. (2016) found a small
positive effect of psychotherapist experience on outcome; and Germer et al. (2022) found no association between experience and outcome (see also Tracey et al., 2014). Thus, the investigation of therapist factors, like ERA, for the selection of very qualified psychotherapists might be relevant.

The second implication is that explicit training of ERA as part of the psychotherapy education might be needed to help trainees develop this ability further. There is first evidence to suggest that trainee psychotherapists’ ERA is associated with therapy outcome and process (Abargil & Tishby, 2021) and if ERA really is an important therapist competency or common factor in psychotherapy, we should find ways to help trainee psychotherapists to improve in this competency. I suggest that standardized computerized training could play a role in psychotherapy education, likely alongside other relevant interventions, like deliberate practice of ERA and other socio-emotional competencies, and other, more psychotherapy-centered interventions, like group seminars about ERA making use of recorded therapy sessions.

Considerations about the undergraduate students’ ERA

Although not focus of this thesis, it is worth discussing that the control group of undergraduate students improved their ERA during one and a half years more strongly than the trainee psychotherapists, contrary to our hypothesis. I can only speculate as to why this was the case. A possible reason for this finding might be that the students “read up” during those months, meaning that they received further education in psychology. The control group was in the beginning of their education whereas the trainee psychotherapists were already in term 7 at the beginning of the study. It is possible that improvements in ERA would not be associated with clinical education but with university education or psychology education in general. Another possible reason might be of developmental nature, as the control group was younger and could have undergone developmental processes relevant for ERA during those months. Yet another reason might be self-selection. Maybe the participants of the control group that came back to the lab were particularly motivated or gifted students that enjoyed taking part in research, which could potentially have led to better results. The fact that we chose undergraduate students as control group in itself can be criticized. Undergraduate students that read classes in psychology might be a biased group to begin with and are not necessarily comparable with the general population, e.g., due to their interest in psychology. They might also already have had classes in psychology that touched upon topics like emotion recognition, empathy or other relevant topics, which could have influenced the results of this study. I would, thus, not be surprised if the differences we found between the two groups at pretest would be more pronounced had we chosen another control group.
Successfully training ERA

To support trainee psychotherapists in improving ERA, my colleagues and I developed two ERA trainings and evaluated them. Both the multimodal ERA training and the micro expression ERA training were found to be effective in a sample of undergraduate students, as well as in a sample of trainee psychotherapists. It is positive that the ERA trainings were effective in two different populations, which indicates that the trainings may indeed be useful for other professions than psychotherapists that could profit from knowledge about and accuracy for non-verbal emotional expressions of others, e.g., pedagogues, social workers, other mental-health professionals or law-enforcement. Since ERA is not yet part of assessment batteries for selecting suitable candidates for the psychotherapy education and since psychotherapy education does not lead to improvements in ERA per se, it is very positive to find that relatively cost-efficient and time-saving interventions can lead to considerable improvements in ERA. Because ERA was found to be relevant for psychotherapy outcome and process (Abargil & Tishby, 2021), psychotherapy programs should consider explicit training procedures for ERA and other relevant therapist competencies and characteristics as part of their education. Nonetheless, replications of this finding and studies investigating the real-life implications of ERA training on therapy outcome and process are very much needed.

Even though (trainee) psychotherapists generally are quite good at recognizing emotions in others, standardized training could at least be relevant for trainees that show relatively low ERA, as we found indication to believe that low ERA is associated with higher gains from training. This is an intuitive finding, as low ERA might leave more room for improvement. The present results propose that the ERA trainings could help compensating for low initial ERA, but more research on this is needed. With more and more research indicating how important individual therapist effects are for psychotherapy outcome and process (Baldwin & Imel, 2013; Heinonen & Nissen-Lie, 2020; Johns et al., 2019; Norcross & Lambert 2006; 2019; Okiishi et al., 2003; Wampold & Imel 2015) and with research showing that psychotherapists already vary considerably in their general efficiency during their education (see, e.g., Edmondstone et al., 2023), it is important to address individual weak spots and work with them early on (or to become more restrictive in selecting suitable candidates). I propose that psychotherapy programs should also consider individualized approaches to training their trainees. It is tremendously important to teach broadly about theories and methods, however, individual trainees might have individual needs when it comes to practical skills. Some might need help with correctly recognizing and working with patients’ emotions, whereas others might need more help with regulating their own emotions during sessions, with expressing warmth or empathy towards patients, with daring to be more confrontative with patients or challenging defenses,
with providing transference interpretations, with learning how to repair alliance ruptures and so on and so forth. Furthermore, depending on whether someone might need support in training multimodal ERA, micro expression ERA or in specific facets of those competencies, e.g., in recognizing emotions based on unimodal displays, ERA training can be individualized. For some psychotherapy modalities, like telehealth or psychoanalysis in a lying-down setting, it might, e.g., be of particular relevance to train reading emotional expressions based on prosody. In other contexts, e.g., in risk assessment or for working with repressed or otherwise defended-against emotions of patients, it could be particularly relevant to train micro expression ERA, and so on. The results of this thesis suggest that multimodal ERA training can be particularly effective for the recognition of unimodal displays. These were the modalities that even showed long-term gains from training. This might be the case because unimodal ERA is more difficult than when several channels of expressions are available, and might, thus, be more responsive to explicit training. That the multimodal ERA training led to long-term effects for unimodal expressions is a very positive result, because it suggests that the training is very potent and potentially of practical relevance. Still, it needs to be noted that these were exploratory analyses with lower item numbers and that the real-life long-term effects on clinical skills still need to be investigated. Further, it needs to be researched how the micro expression training could be modified to lead to long-term effects as well.

That the trainings did not produce transfer effects to the measure for verbal and non-verbal ERA in medical contexts suggests that ERA training needs to be quite explicit to be effective. Tailormade training of specific competencies that are relevant to individual trainee psychotherapists might have a place in psychotherapy education in the future. Both trainings were rather short, compared to physical interventions or cognitive trainings, leading to the question whether more intensive trainings could have led to stronger long-term effects or possibly transfer effects. However, I am doubtful that more or longer training sessions alone would have led to better retrieval at follow-up. Instead of providing more of the same training, I suggest that the standardized computerized trainings should be part of larger training approaches for improving socio-emotional and relationship skills to reliably lead to long-term and practically meaningful improvements. Thus, training should not only be specific and individually tailor, but also, in other aspects, broad enough. Standardized ERA training should be combined with more individualized and naturalistic training components, like deliberate practice of ERA in role-plays or in video-supported supervision that would make training more ecologically valid and more meaningful to the trainees.
Methodological considerations

There also needs to be a methodological discussion of the studies in the present thesis. Generally, the within-between group design of all three studies is positive as it allows for higher statistical power, and information about long-term development of ERA and long-term effects of the trainings. It allows to take into account individual ERA baseline values (intercepts) and individual change slopes instead of only focusing on cross-sectional group differences.

To my knowledge, this is the first study that includes a long-term follow-up measurement for investigating the long-term effectiveness of ERA training for mental health professionals and also the first study that includes training of unimodal auditory ERA. Also, to my knowledge, this is the first study that follows trainee psychotherapists’ ERA during a longer time span. The only other known study (Machado et al., 1999) about how experience affects ERA was cross-sectional. Further, the investigation of two different, but related ERA facets – multimodal ERA and micro expression ERA – is a methodological strength of the present thesis, since it provides a more detailed picture of ERA changes during psychotherapy education and the trainability of ERA, and largely confirms that different facets of ERA don’t improve in a meaningful fashion during the psychotherapy education and, at the same time, that ERA generally can be successfully trained, even though the recognition rates for the two tasks vary. Furthermore, it is worth noting that the training studies were randomized and controlled, single-blind (Study II) and double-blind (Study III) studies that allow drawing causal inferences.

Study design and statistical methods

Obvious issues that need to be discussed are sample size, drop-outs, and statistical power. In Studies I and III, we did not specify the sample size in advance, but opted to include as many volunteers as possible. This was, partly, a pragmatic choice. The sample size of Study II was chosen based on the expected sample size of Study III. Because we were interested in a very specific group of participants – trainee psychotherapists about to start with their specialized psychotherapy training – and since our research group had funding for three enrollment periods, we were dependent on how many trainees would be willing to participate. We compensated them for their participation (e.g., with gift cards, sandwiches and course credits) and managed to include relatively many participants in all three studies, compared to other studies in the field. We also tried our best to get the participants to come back to the lab after one to one and a half years. However, not all participants chose to do that. Particularly Study I had considerable dropout, which limits the interpretability of the result that psychotherapy education does not lead to significant improvements in ERA. Although some dropout is to be expected in longitudinal research, Study I had a dropout rate of 53% in the 1.5-year follow-up. The
dropout might have been related to the participants’ age with younger participants being more likely to drop out. The dropout was not related to individual ERA scores. Because of the dropout, we ran the mixed multilevel analyses with maximum likelihood estimation to handle missing data, and performed additional sensitivity analyses in which we dropped all the missing participants from the analyses altogether. Whereas the results were not identical – the analysis with maximum likelihood estimation found that the control group had significantly steeper change slopes than the trainee psychotherapists whereas the sensitivity analyses did not find any interaction effects – the conclusion was the same. The hypothesis that trainee psychotherapists would improve more strongly in ERA due to their education was not supported in either analysis. Generally, too low sample size can be associated with low statistical power and, thus, lead to an increased probability of false negative results. Nonetheless, the maximum likelihood estimation results and the descriptive data in Study I speak against this being the case. The problem of regression to the mean that is known to occur in longitudinal research without randomized groups or correction for covariates also needs to be mentioned. The phenomenon of regression to the mean, in this context, means that very high or very low ERA values could become more moderate in the next measurement. The idea is that extreme values oftentimes contain higher degrees of measurement error, which might lead to less extreme values under other circumstances. We cannot preclude that this phenomenon could have influenced the results. It would have been beneficial to have more than two measurement time points, possibly a multiple baseline design, or to correct for it in the statistical analyses. On the other hand, we could also assume that regression to the mean might have influenced both groups equally.

In Study II, we used a rather unconventional statistical analysis – the Aligned Rank Transform (Wobbrock et al., 2011) for mixed factorial designs – because assumptions for parametric ANOVA were not all consistently fulfilled (e.g., non-normally distributed residuals and/or heterogeneity of variance in some cases). This might limit the comparability with other studies using different statistical methods and did not allow for within-between group effect size estimates. However, the results of Study III (using mixed multilevel modeling) support the findings of Study II. In hindsight, using a Bayesian approach that is not assuming a Gaussian distribution would have been a relevant alternative. Also the other two studies could have been analyzed with Bayesian statistics. In a pervious draft of Study I, we also performed supplementary Bayesian ANOVAs that confirmed the result that psychotherapy education does not lead to significant improvements in ERA, but the analyses did not make it into the final draft. The data for multimodal ERA suggested that the null-model (no effect whatsoever) was more likely than the existence of effects of time (beginning of therapy education to end of psychotherapy education), group (trainee psychotherapists vs. control group) or an interaction. The
data for micro expression recognition suggested overwhelming evidence for the main effect of time relative to the null-model, but did not support an effect of group or an interaction.

Reliability issues
Another relevant issue is the reliability of the ERA tasks. The MICRO showed acceptable to good internal consistency according to Cronbach’s alpha throughout the three studies. The ERAM, on the other hand, showed lower Cronbach’s alpha values at several occasions in the three studies. This suggests some instability of the measurement instrument and limits the trustworthiness and interpretability of the results for multimodal ERA. Establishing sufficient reliability of measures is evidently important, also in emotion recognition accuracy research (see Olderbak et al., 2021). However, receiving high Cronbach’s alpha values for a task including twelve emotion categories and three modalities might be challenging. One reason for that could be that the ERAM includes items with varying intensity and difficulty levels (see Bänziger et al., 2012; Laukka et al., 2021) and that we cannot assume that all items stand in equal relationship to one another (e.g., anger and irritation, two similar emotions varying in intensity, are supposed to have a stronger relationship with each other than, e.g., anger and joy) and this might affect the internal consistency of the test overall. According to Olderbak et al. (2021), individual biases in perceiving certain emotions could also violate the assumption of independence of items needed for alpha estimates. They suggest that the omega coefficient (McDonald, 1999) might be preferable to alpha, if only one measure of reliability is used (even though they suggest to report several whenever possible). Omega values are based on factor-analytic models and can be interpreted as measures for internal consistency of a test as well as for dimensionality or internal structure (Flora, 2020; Olderbak et al., 2021). The Omega total estimates (see McDonald, 1999; McNeish, 2018; Revelle, 2009; Revelle & Zinbarg, 2009) in Study I suggest that the proportion of variance explained by ERA as general factor and the twelve latent factors (the emotion subscales) was satisfactory and that the reliability of the ERAM according to omega total coefficient is still acceptable. Furthermore, the test-retest correlations of the ERAM were of moderate size, as would be expected in long-term measurements and cases of ERA improvement. In two evaluation studies (Laukka et al., 2021), the ERAM did not show reliability issues and displayed relevant correlations with related concepts, confirming its criterion validity.

In a previous draft of Study I, we tried to modify the ERAM by dropping items with low item-total correlations from the test, but we did not find this to lead to considerable improvement of the alpha values. Thus, we decided against using a decimated version of the ERAM as this would have led to uneven modality/valence/arousal/emotion scales. Still, we reran the analyses of the previous draft with the modified ERAM and found the same results. All
in all, the question surrounding the reliability of the ERAM is, in my view, not conclusively answered. The use of the ERAM in the three studies was decided based on previous encouraging research about the ERAM (see Laukka et al., 2021) and it is unfortunate that the ERAM does not show consistently good reliability in the present thesis. I argue that the results regarding multimodal ERA and the trainability of this ERA facets still can be seen as valid, but would like to note that future research has to investigate the ERAM and its reliability further.

Specifics of the tasks and trainings

Further, the specifics of the ERA tasks and trainings can be discussed. One relevant issue is the use of standardized displays of emotional expressions by actors and trained lay-people for creating the stimuli. Experts (Scherer & Bänziger, 2010) suggest that the use of actors in creating emotional expressions is a valid approach and has its place in ERA research, among other methods. Still, there is always a shred of doubt about whether staged emotional expressions actually are recognizable as “real” emotions. Recognition rates most certainly depend on how “well”, how stereotypical and how intensive an emotion is displayed by the expresser (see, e.g., Khosdelazad et al., 2020). In real life interactions, e.g., in psychotherapy, emotional expressions might be much less stereotypical or obvious. Furthermore, patient might display mixed, conflicting, blended or otherwise unclear emotional expressions. The two ERA measures incorporate stimuli that showed above chance levels of recognition in evaluations. The MICRO includes expressions according to Ekman et al.’s (2002) *Facial Action Coding System* that is systematically describing which facial muscles are involved in facial emotional expressions. The ERAM made use of items with varying levels of (above chance) inter-rater agreement and different levels of difficulty to not make the task too easy. The main advantage of standardized tasks is that they allow for an “objective” assessment of an ability, which improves the comparability of ERA rates. This is why we chose this kind of measurement. Those tasks are also more resource efficient.

This leads us, however, to another issue: the items were not tailor-made for the psychotherapy context. Because we also included participant groups of undergraduate students that did not have contact with psychotherapy education, it was more feasible to use standard ERA tasks. However, for the trainee psychotherapists it would have been relevant to (also) include a more specified non-verbal measure and training for ERA in psychotherapy situations to make the tasks and trainings more meaningful for them. This would likely also have increased the external validity of the findings. We included the PECT for that reason, but it is a test for both verbal and non-verbal ERA and it includes typical medical interactions, in other words, it was not a perfect measure for this purpose either. Further, the forced-choice nature of ERA tasks and trainings could potentially lead to response bias. Participants could, e.g., guess
what the closest answer might be and choose it from the list, even in cases where they don’t really recognize the emotion. There is also the well-known problem of ceiling effects for the emotion happiness, as it oftentimes is the only unambiguously positive emotion in basic emotion tasks, like the MICRO (see, e.g., Kessels et al., 2014). It is generally recommended to take into account false alarm rates in those kinds of tasks (Griffiths et al., 2015). To circumvent individual response biases, we chose to use Wagner’s (1993) unbiased hitrate for the analyses. The unbiased hitrate is correcting for how often an answer was given incorrectly (not only how often it was given correctly). For example, if someone used the emotion label “sadness” very often, also for other emotions, then the recognition rate for this emotion is not only based on the true positives for sadness (average correct) but also on how often the label was incorrectly applied to other emotional expressions. The recognition rate for sadness might be near perfect, but only because the participant applied it many times. The unbiased hitrate is a way of controlling for response bias, namely for how often an emotion category was used incorrectly (and not only the average number of correct answers for a certain emotion). Taking into account false alarms has been shown to be relevant in ERA training studies (see e.g., Griffiths et al., 2015). Another issue might be the high number of emotions, and also similar emotions, of the ERAM. It is possible, that the results of the studies could have been driven by specific emotions that are either very easily recognized/trained or very difficult to recognize/train. For example, had the ERAM included less items, or less similar items (like anger and irritation), the recognition rates or training success might have been stronger. This might also be true for the internal consistency of the ERAM. Less similar items might have led to less reliability issues.

Psychotherapeutic approach

Another methodological point is the use of PDT and CBT trainees in the same study. We chose to include both psychotherapy approaches taught at Stockholm University on the one hand for pragmatic reasons. By including both approaches, we were more likely to receive an adequate sample size. Luckily, the analyses of group differences between PDT and CBT trainees suggest that they showed comparable ERA and that it was unproblematic to summarize them under the term “trainee psychotherapists”. But even beyond pragmatism does it make sense to investigate them as one group. I consider ERA to be a therapist competency that is relevant for all kinds of psychotherapy, as a potential common factor in psychotherapy. Theoretically, both approaches work with patients’ emotions and emotional expressions, although in different ways, and being able to read patients’ emotional expressions should be important in both (or all) psychotherapies, as we know about therapist characteristics like Rogers’ three therapist attributes empathy, positive regard and genuineness (see Norcross & Lambert, 2019). With that being said, future studies
should still try to investigate whether there are some specificities for different approaches, e.g., when looking more closely into the work with emotion recognition in psychotherapy and interpersonal dynamics, using bigger samples or case studies. The choice to include trainee psychotherapists instead of fully experienced psychotherapists also had pragmatic as well as conceptual reasons. A lot of psychotherapy research is conducted at university clinics, as the recruitment of participants and implementation of studies tends to be less complicated. But it also makes sense to investigate trainees because meta-analyses (Bohart et al., 2002; Elliott et al., 2011) showed that the relationship between therapist empathy and therapeutic outcome is moderated by therapist experience, in so far that the relationship is stronger for less experienced therapists. Empathy and emotion recognition might be effective tools on which psychotherapists can “fall back on” in case of less theoretical and methodological tools.

Summary of strengths and limitations

The present thesis has several strengths and limitations, most of which I already discussed above. Strengths are, e.g., the within-between design of the studies and that the participants were randomized into experimental and control groups. This provides the opportunity to investigate changes in ERA and compare those changes between the groups to try to draw conclusions and make causal inferences. The thesis also contains the first long-term follow-up investigation of the effectiveness of ERA training for mental health professionals. This is of special relevance as it is not unplausible that repeated ERA training might lead to effects in the short-term, but it is much less certain that those effects last. Most previous studies conducted the posttest directly after the training. Further, it is positive that we used two different, but related ERA measures, as this can provide us with a more differentiated picture about ERA changes due to education and training, as well as strengthens the findings of this thesis, since the multimodal and micro expression ERA changes show quite similar patterns, at least in the short term. It is also noteworthy that this has been the first time that multimodal and unimodal ERA was measured and trained in the population of trainee psychotherapists and other mental health professionals. Especially the recognition accuracy for prosodic expressions is an understudied field. This is why it is encouraging to see that auditory ERA indeed can be trained, also in psychotherapists.

But the present thesis also has several limitations. One is the already mentioned high dropout in Study I, which somewhat limits the interpretability of the result that psychotherapy education does not lead to ERA improvements. Related to that, it has to be mentioned that we used a sample of convenience and did not define the sample size in advance, but relied on including as many participants as possible during three terms of data collection (for Study I and
Although the sample sizes of the studies were relatively large compared to other research in the field, it was, e.g., not possible to select only PDT or only CBT students for analyses, as to not diminish the sample size and lose statistical power. Another limitation is the reliability of the ERAM that is not fully ensured, even though I argue that its shortcomings can be understood in the light of the specifics of the task (e.g., many scales and items with varying relationship to one another). A further methodological limitation might be that we did not use a naturalistic measure for ERA that was tailor-made for trainee psychotherapists, which leaves the question unanswered, whether some aspects of ERA that are related to psychotherapy might have changed during the 1.5 years of psychotherapy education, but simply could not be measured with the standardized tasks that we used. In other words, one can question the validity of the tasks, as it is unclear whether the standardized computerized ERA tasks actually measure practically or clinically meaningful emotion recognition competencies. Further, there was a coding error for anger in the micro expression training and, thus, this emotion was excluded from the analyses or analyzed separately in the training studies. Subsequently, it is unclear in how far anger micro expression recognition is trainable using the micro expression training. Further, the active control training only contained immediate feedback in the beginning of the task, thus, is not completely comparable to the other two training groups. In fact, as we report in Study III, most participants that were part of the control group had also guessed that they did not receive the actual training. This suggests that we cannot rule out motivational factors in the lower ERA performance of the control group. The limited research basis for the assumption that ERA is a relevant therapist competency to possess, is another weakness of the study. However, we hope to solve this in the future by looking into the therapy outcome data for the patients of the participants of Study I and III. Lastly, the problem of generalizability of the results needs to be addressed. The similar findings of the two training studies propose that the ERA trainings might be useful even for other populations. However, the generalizability of the finding that psychotherapy education does not lead to improved ERA is less certain. It is possible that other, more specified psychotherapy approaches might lead to ERA changes, or that ERA training might influence more experienced psychotherapists differently.

Open questions and future directions

This thesis was part of two larger research projects and several studies will still be published. I will investigate the relationship between trainee psychotherapists’ ERA (in the beginning and the end of psychotherapy education) and psychotherapy outcome for the sample of trainee psychotherapists that were part of Study I. For that, we collected the psychotherapy outcome data
(pre–post therapy) of the consenting patients of the trainee psychotherapists. Psychotherapy outcome data and measures of the therapeutic alliance were also collected for the consenting patients of participants of Study III. This provides the opportunity to also investigate how training ERA might influence therapy outcome and process. These studies are important because there is, so far, only one study that experimentally studies the influence of ERA on psychotherapy process and outcome (Abargil & Tishby, 2021) and there is need for replications before we can be more certain that (trainee) psychotherapists’ ERA is of real importance in psychotherapy. And further, that ERA training (potentially) leads to real life effects for patients. Further, for the sample of Study III, I also collected the participants’ facial expressions while conducting the ERA tasks. The facial patterns while watching, recognizing and matching emotional expressions could provide insights about ERA and the process of learning ERA. Further, a colleague also conducted interviews with the participants of Study I about emotions in psychotherapy, which provides a qualitative approach to understanding how trainee psychotherapists experience working with emotions and emotional expressions in psychotherapy. Another study about nonverbal behavioral dynamics in psychotherapy is in preparation. In this study, we will investigate the non-verbal interplay of trainee psychotherapists, experienced therapists and experienced psychotherapists working online and their patients using machine learning and time-series analyses of the interactions. The work with artificial intelligence provides the opportunity to analyze the non-verbal interplay in psychotherapy much faster and in more detail than when using manual coding procedures.

Beyond those already conducted or planned studies, there is the general need for conceptual replications. Those studies need to find better ways of preventing dropouts and could explore different versions of the trainings (e.g., including only specific items). Firstly, there need to be more studies establishing whether there is a meaningful contribution of trainee psychotherapists and psychotherapists ERA to therapy process and outcome. Secondly, there is the need for replications of the results presented in this thesis and modifications of the studies. It remains to be explored, e.g., why trainee psychotherapists are showing higher ERA than undergraduate students and this can only be understood using even longer studies, preferably assessing ERA already at the beginning of the psychologist program or earlier, and taking into account variables like academic success, motivation and personality traits. Further, although the present thesis suggests that trainee psychotherapists’ ERA does not meaningfully improve during the course of standard PDT and CBT psychotherapy education, it is possible that other aspects of ERA than those assessed in the present studies could be improving. This idea is supported by the results of Machado et al., (1999) that used modified video-clips of therapy sessions as measure of ERA. Other ERA measures that are tailor-made for the psychotherapy context or ERA measures with less reliability issues (than the ERAM)
should be used in replications. The same might be true for ERA training studies. Although we do not yet know whether ERA training for trainee psychotherapists leads to mental health improvements in their patients, it is not unlikely that some additional training components that focus on ERA in the therapy room, like the use of video-feedback, deliberate practice, and other interventions focusing on patients’ emotional expressions might be needed in addition to computerized training. This should be investigated in future studies. Thirdly, the ERA trainings should be extended to other populations that might profit from them, e.g., to other professions that rely on emotional and empathic exchanges, like teachers, nurses, physicians, or law-enforcement.

Further, there is a multitude of possible interesting studies based on the results of this thesis that should be conducted. It is, for example, possible that certain emotions bear special relevance for the psychotherapy context. I’m intuitively thinking about emotions like shame, guilt, irritation, sadness, anger or fear, but also other emotions could be specifically relevant for certain therapy formats or patients. So, future research could focus more on specific emotions. Also, in real life encounters, e.g., in psychotherapy, emotions are not always sequential and patients often describe having different, possibly conflicting emotions at the same time. Thus, it might be relevant to investigate mixed, blended or conflicting emotions in the context of psychotherapy. Further, from the results presented in this thesis, we know that emotions communicated via prosody are more difficult to recognize than audio-visual expressions or even visual expressions. Future research should investigate how the ERA for different modalities varies in different psychotherapeutic approaches and settings that work with unimodal displays of emotion, like in tele-health or online settings, or in classical psychoanalysis in a laying-down setting. Trainings for certain modalities for certain settings might be in order.

Furthermore, the question arises whether it really is important for psychotherapists to be able to recognize which distinct emotion someone is (briefly) displaying or whether it might be enough for a psychotherapist to be able to recognize that there were some emotional shifts potentially happening within the patient and to use this observation as a starting point for joint exploration of emotions. It might also be interesting to explore whether there is a difference between (trainee) psychotherapists’ immediate ERA (as assessed via standardized tasks including snippets of emotional expressions) and retrospective ERA or longer recognition processes that happen during a whole therapy session, possibly even including the combination of verbal and nonverbal content. Patients are no strangers to their psychotherapists and, thus, it might be relevant to also investigate ERA for personally-known individuals.

As a psychodynamic therapist myself, I would love to investigate how (trainee) psychotherapists’ ERA is associated with psychodynamic concepts, as I consider this kind of interdisciplinary research to be of tremendous importance, both for psychodynamic therapy and theory, as well as for broadening and enriching basic psychological research with psychodynamic thinking.
(see also Hau, 2015). In this context, I would, e.g., be interested in the relationship between ERA and transference, countertransference and mentalizing capacities. We know that transference happens in everyday life (see, e.g., Andersen & Berk, 1998; Chen & Andersen, 1999) and psychotherapy (see, e.g., Gelso et al., 1997; Gelso & Bhatia, 2012), but we know very little about how transference dynamics evolve and how we communicate nonverbally via transference and countertransference. It is not hard to imagine that therapists that misinterpret patients’ emotional expressions might have distorted perceptions of the patient’s transference and, thus, possibly distorted countertransference reactions (possibly related to own projections and fantasies in response to supposed expressions, i.e., “neurotic” countertransference reactions). In future studies, I would like to investigate psychotherapists’ ERA and transference as measured by the therapist, the patient and observers, as well as the therapists’ countertransference. Further, I also would like to include a measure for the trainee psychotherapists’ mentalizing capacities in future studies, especially affective mentalizing and mentalizing others. To be able to read external cues of other people’s mental states – e.g., to recognize their emotional expressions – is needed to mentalize others’ inner states accurately (Frith & Frith, 2006; Mitchell & Phillips, 2015). Neuroscientific research also suggests a link between mentalizing about others’ emotions and empathy (Hooker et al., 2008). Thus, it would be valuable to investigate the link between the trainees’ ERA and different facets of mentalizing. Further, it would also be interesting to investigate the link between therapists’ emotion regulation or mentalized affectivity and embodied mentalizing (Jurist, 2018) and ERA in psychotherapy.

Finally, nonverbal and verbal relational dynamics in psychotherapy should ultimately be investigated as dyadic processes (Atzil-Slonim & Tschacher, 2020). There is research that suggests that bio-behavioral relational interplays become more advanced the closer the attachment bond is between subjects (see, e.g., Feldman, 2017). While, e.g., behavioral coordination of culture-specific display rules already occurs with strangers, for actually synchronized behavioral (e.g., facial expressions, body-language, prosody) and psychophysiological (e.g., heart rate) reactions, there needs to be a closer human attachment (like in parent-child interactions or romantic partnerships). Emotion recognition research about psychotherapeutic encounters should ultimately investigate both parts of the interaction and assess nonverbal synchronies in psychotherapy, and possible associations with psychotherapists’ and patients’ ERA.
Conclusion and clinical relevance

The present thesis investigated trainee psychotherapists’ ERA before and after psychotherapy education and how this ability can be trained using standardized computerized training procedures. I found that trainee psychotherapists’ multimodal and micro expression ERA was greater than a sample of undergraduate students’ ERA. However, this thesis suggests that trainee psychotherapists’ ERA does not necessarily improve due to standard PDT or CBT education. Still, this thesis also suggests that trainee psychotherapists’ (and undergraduate students’) multimodal and micro expression ERA can be successfully trained with standardized computerized trainings. These findings are contributing to the research about relevant therapist characteristics and competencies and how they potentially change (or not change) with experience and education. Further, they give practical guidelines for how ERA could be trained in resource-efficient ways, to support trainee psychotherapists in becoming more perceptive of their patient’s emotional expressions, and thus, hopefully, becoming better psychotherapists. Improving trainee psychotherapists’ ERA could even have ethical implications, not only for the trainee psychotherapists that might feel more enabled to establish a durable relationship with patients and help them in more efficient and empathic ways, but potentially also for their patients. Improved ERA of trainee psychotherapists could potentially even contribute to better and safer care for patients. However, this notion still needs to be explored systematically in future studies.
References


74


Constantino, M. J., Boswell, J. F., Coyne, A. E., Kraus, D. R., & Castonguay, L. G. (2017). Who works for whom and why? Integrating therapist effects analysis into psychotherapy outcome and process research. In L. G. Castonguay & C. E. Hill (Eds.), *How and why are some therapists better than others?: Understanding therapist effects* (pp. 55–


