RICERCHE DI PSICOLOGIA

PSYCHOLOGICAL RESEARCH JOURNAL

NUOVA SERIE - ANNO XLVII
N. 3, 2024

Trimestrale fondato da Marcello Cesa-Bianchi

RICERCHE DI PSICOLOGIA

Trimestrale fondato da Marcello Cesa-Bianchi

Fondatore: Marcello Cesa-Bianchi.

Comitato Direttivo: Alessandro Antonietti (Università Cattolica del Sacro Cuore, Milano), Paolo Inghilleri (Università degli Studi di Milano), Antonella Marchetti (Università Cattolica del Sacro Cuore, Milano).

Segreteria di redazione: Anna Mottini (Università Cattolica del Sacro Cuore, Milano).

Comitato Scientifico: Susana Alves (Università di Roma La Sapienza), Ilaria Castelli (Università di Bergamo), Barbara Colombo (Champlain College, Burlington), Amedeo D'Angiulli (Carleton University, Ottawa), Gaspare Galati (Università di Roma La Sapienza), Omar Gelo (Università del Salento), Michaela Gummerum (University of Warwick), Antonio Iannaccone (Université de Neuchâtel), Beatrice Ligorio (Università di Bari Aldo Moro), Todd Lubart (Universitè Paris Descarts), Francesca Pazzaglia (Università di Padova), Egidio Robusto (Università di Padova), Mario Rossi-Monti (Università di Urbino Carlo Bo), Ai-Girl Tan (Nanyang Technological University, Singapore), Elena Vegni (Università degli Studi di Milano), Adriano Zamperini (Università di Padova).
Comitato Onorario: Piergiorgio Argentero (Università di Pavia), Dora Capozza (Università di Padova), Carlo Cipolli (Università di Bologna), Cesare Cornoldi (Università di Padova), James C. Coyne (University of Pennsylvania), Dario Grossi (Università della Campania Luigi Van- vitelli), Riccardo Luccio (Università di Trieste), Barbara Ongari (Università di Trento), Anne- Nelly Perret-Clermont (Université de Neuchâtel), Louis Ploton (Université Lumière Lyon 2), Marco Poli (Università di Milano), Paolo Renzi (Università di Roma La Sapienza), Ian Robertson (Trinity College, Dublin), Wolfgang Schnotz (Universität Koblenz-Landau).

Redazione: Dipartimento di Psicologia, Università Cattolica del Sacro Cuore, Largo Gemelli 1, 20123 Milano – Tel. 02 7234 2284; *e-mail*: ricerchedipsicologia@unicatt.it.

Amministrazione
FrancoAngeli srl, viale Monza 106, 20127 Milano – Tel. 02 2837141
Autorizzazione Tribunale di Milano n. 128 del 30 marzo 1976 – Direttore responsabile: dr. Stefano Angeli – Trimestrale – Copyright © 2025 by FrancoAngeli srl.
Pubblicato con licenza Creative Commons Attribuzione - Non Commerciale - Non opere derivate 4.0 Internazionale (CC-BY-NC-ND 4.0) L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito https://creativecommons.org/licenses/by-nc-nd/4.0/deed.it

III trimestre 2024

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

INDICE N.3, 2024

Early Network-based Action Against Violent Behaviors to Leverage victim Empowerment: Results from an action-research plan, di Francesca Ranghetti, Sarah Miragoli, Paolo Giulini, Benedetta Rho, Laura Emiletti, Marina Contino, Elvira Tassone, Umberto Battaglia, Elisabetta Corbucci, Silvia Merli, Susanna Murru e Luca Milani	Pag.	7
Cracking the Code: Examining Psychometric Rigor of the Provider Decision Process Assessment Instrument (PDPAI) among Residents' Trainees and Expert Physicians, di Sana Younas e Saeeda Khanum	"	31
Intelligenza emotiva dell'insegnante e relazione insegnante-alunno nella scuola primaria. Uno studio esplorativo, di Giordana Szpunar, Chiara Mannola, Eleonora Cannoni, Madalina Grigore e Anna di Norcia	"	57
The relationship between theory of mind and emotion regulation in the preschool years. The role of language and gender, di Daniela Bulgarelli, Maria Nives Sala, Beatrice Ragaglia e Paola Molina	"	81
Teaching Clinical Reasoning and Critical Thinking in Psychology Graduate Courses: A Systematic Review, di Tugce Elbir, Catia Cammeo, Francesca Cavallini, Antonella Marchetti e Davide Massaro	"	105

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

TABLE OF CONTENTS N.3, 2024

Early Network-based Action Against Violent Behaviors to Leverage victim Empowerment: Results from an action-research plan, di Francesca Ranghetti, Sarah Miragoli, Paolo Giulini, Benedetta Rho, Laura Emiletti, Marina Contino, Elvira Tassone, Umberto Battaglia, Elisabetta Corbucci, Silvia Merli, Susanna Murru e Luca Milani — Pag. 7

Cracking the Code: Examining Psychometric Rigor of the Provider Decision Process Assessment Instrument (PDPAI) among Residents' Trainees and Expert Physicians, di Sana Younas e Saeeda Khanum — " 31

Teacher emotional intelligence and teacher-student relationship in primary school. An exploratory study, di Giordana Szpunar, Chiara Mannola, Eleonora Cannoni, Madalina Grigore e Anna di Norcia — " 57

The relationship between theory of mind and emotion regulation in the preschool years. The role of language and gender, di Daniela Bulgarelli, Maria Nives Sala, Beatrice Ragaglia e Paola Molina — " 81

Teaching Clinical Reasoning and Critical Thinking in Psychology Graduate Courses: A Systematic Review, di Tugce Elbir, Catia Cammeo, Francesca Cavallini, Antonella Marchetti e Davide Massaro — " 105

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Early Network-based Action Against Violent Behaviors to Leverage victim Empowerment: Results from an action-research plan

Francesca Ranghetti*, Sarah Miragoli*, Paolo Giulini°, Benedetta Rho°, Laura Emiletti°, Marina Contino^, Elvira Tassone^, Umberto Battaglia**, Elisabetta Corbucci°°, Silvia Merli§, Susanna Murru^^, Luca Milani*

* CRIdee - Department of Psychology,
Università Cattolica del Sacro Cuore, Milan, Italy;
e-mail: francesca.ranghetti@unicatt.it;
e-mail: sarah.miragoli@unicatt.it;
e-mail: luca.milani@unicatt.it.
° CIPM soc. coop. soc.,
Italian Center for the promotion of Mediation, Milan, Italy;
e-mail: paolo.gulielmo.giuliniGgmail.com;
e-mail: benedetta.rho@gmail.com;
e-mail: laura.emiletti@gmail.com.
^ Anti-Crime Central Directorate of the State Police, Rome, Italy;
e-mail: marina.contino@poliziadistato.it;
e-mail: elvira.tassona@poliziadistato.it.
** Associazione Ares APS,
Violence Offender Treatment Center, Bassano del Grappa, Italy;
e-mail: umberto.battaglia@gmail.com.
°° CIPM Liguria aps,
Italian Center for the promotion of Mediation, Genova, Italy;
e-mail: lisacorbucci@yahoo.it.
^^ CIPM Sardegna ODV,
Italian Center for the promotion of Mediation and Restorative Justice, Cagliari, Italy;
e-mail:susamurru@gmail.com.
§ CIPM Emilia Impresa Sociale,
Italian Center for the promotion of Mediation, Piacenza, Italy;
e-mail: merli.silvia.psi@gmail.com.

Submitted: 09.08.2023 - **Accepted:** 15.07.2024

Available online: 23.12.2024

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Acknowledgments

No conflicts of interests.

Funding Statement

The research-intervention referred to in this article is funded by the EU REC-RDAP-GBV-AG-2020 “ENABLE” call.

Ethics Statement

Since all of the data was collected without any contact by the research team with the participants, and no personal or sensitive data was collected, an Ethical vetting was not required.

Abstract

The ENABLE project, which aims to implement the ZEUS protocol at the national level, aims to activate specific interventions to prevent violence against women by going directly to individuals who have been admonished for domestic violence or stalking, thus implementing a system of protection towards victims. A second objective concerns the implementation of a training intervention aimed at project workers and law enforcement officers. The topic of gender-based violence, identification and management of situations at risk of escalation through the construction of shared operational practices was the subject of the training. Following the trainings, a quantitative survey was administered to law enforcement officers (both training participants and non-trainees) to assess their knowledge of the instrument of warning, referral to treatment and specific forms of gender-based victimization, use of referrals to treatment, and knowledge of gender-specific forms of victimization. The results showed significant differences between participants versus non-participants to ENABLE training, with respect to intervention effectiveness, awareness and knowledge related to the training. With regard to the admonished and petitioning parties, two different surveys were administered aimed at investigate, at the end of intake, clinical and individual adjustment variables (in terms of depression and posttraumatic symptoms, coping strategies, risk of victimization, and adverse childhood experiences). The results showed significant differences between the two groups (admonished versus instant parties), particularly with regard to the depressive and post-traumatic risk variables, with the instant parties being more likely to be exposed to this condition than the admonished.

Keywords: PV, treatment of offenders, admonishing, victims protection

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Introduction

The ENABLE (Early Network-based Action against abusive Behaviours to Leverage victim Empowerment) project, funded on the REC-RDAP-GBV-AG-2020 "ENABLE" call, is a research-intervention project aimed at preventing gender-based violence, protecting victims and activating integrated intervention models, with particular reference to the social perception of the phenomenon of domestic violence and stalking, and the taking charge of warning situations by the Questore.

The project which aims to implement the 2018 ZEUS protocol at the national level, a memorandum of understanding between the Milan Police Headquarters (Central Anticrime Service, Ministry of the Interior) and the Italian Center for the Promotion of Mediation (CIPM), for the effective implementation of Law 38/2009 and 119/2013, concerning the crimes of persecutory acts – stalking – and domestic violence.

In Italy there is the device of the Questore's "warning", that is, a formal warning issued by the Questore to the reported person to refrain from committing further acts of harassment and/or violence and understand its social disvalue. Contextually, the warning allows the Questore to take measures such as the withdrawal of weapons and – in case of reiteration of conduct, ex officio prosecution – to prosecute ex officio.

The practice defined by the Zeus protocol stipulates that concurrently with the warning, subjects are sent to the CIPM for clinical-criminological treatment.

Likewise, victims are given information and referrals regarding territorial support and advocacy services such as anti-violence centers. They are also asked for their willingness to be contacted by the CIPM for more accurate situation monitoring and risk assessment.

It is important to clarify that a "warning" does not mean a complaint against the perpetrator of violence, but is a "*measure designed to 'recall' or warn' a person who engages in persecutory and/or violent acts, which do not yet constitute a crime, so that he or she will stop his or her harassing activity before it escalates to a worsening that makes the activation of a criminal trial inevitable"* (ARES Association APS, 2003). This prevention perspective is also concerned with the possible escalation of the frequency, intensity, and harmfulness of victim-protection conduct.

The following participated as Project partners: the CIPM (soc. coop. soc. as lead partner), C.I.P.M. Liguria APS, The CIPM Emilia Impresa Sociale, C.I.P.M. Sardegna, Associazione ARES APS, and the Catholic University of the Sacred Heart of Milan as scientific referent.

Associate partner was the Ministry of the Interior, Servizio Centrale Anticrimine (SCA). With the ENABLE project it was possible to replicate

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

in 10 Italian cities the model of the ZEUS Protocol, which had already been initiated in the city of Milan. The goal of the project was the implementation of good practice that allowed: the early interruption of violent behavior and the reduction of recidivism, the protection of victims and the knowledge of territorial support resources, the definition of a multiagency model, and the awareness of stakeholders and the community.

The Zeus protocol aims at the awareness and empowerment of the admonished through reframing their acts and gaining awareness of the effects of their violent behavior. In particular, potential abusers are offered the treatment opportunity to question their relational patterns and views on gender relations (cf. Garbarino & Giulini, 2022, 2020).

In addition, a number of training days were conducted, addressed to the Forces dell and project operators in which some pivotal issues of interventions in situations of gender-based violence were addressed and delving into the operational practices of the ZEUS Protocol, such as:

- The legal context of the warning and the application of the Zeus Protocol;
- The opportunities and methods for sending to treatment;
- The techniques for collecting the testimony of the petitioning party for the issuance of warnings;
- Reception and interviews with the petitioning parties and referral to the appropriate territorial services;
- The recognition of high-risk individuals, through ll recognition of situations at risk of escalation or reiteration;

The construction of territorial networks.

In addition, 6 more trainings have been conducted at police headquarters in some of the cities involved in the project in a logic of cascade training (train the trainer) and training packs that allow people to take advantage of the training content

The following paragraphs will detail the results (quantitative and qualitative) collected within the ENABLE Project, with respect to the Law Enforcement Agencies, the admonished, and the petitioning parties.

Gender-based violence

In 2002, the Council of Europe defined violence against women as "*any act of violence based on sexual affiliation that results or is likely to result in harm or suffering of a physical, sexual or psychological nature to women, who are targeted by it"* (Recommendation REC (2002) of the Council of Ministers of member states on the protection of women from violence). In June 2013, Italy ratified the Istanbul Convention, which

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

identifies violence against women as "*a violation of human rights*" and includes any act or threat within the scope of gender-based violence that causes or is likely to cause physical, psychological, sexual or economic harm or suffering in both public and private life. With this document, the focus is also placed on violence in private life, and thus violence within the home is placed in a perspective according to which it is not merely a private matter, but one of public significance. With the coining of the term "*intimate partner violence*", or *intimate partner violence*, we go on to specify and define that form of violence against women, committed in the couple relationship system, by current or former partners: this type of violence against women is the most frequent within the domestic walls (Bonura, 2018).

Among the many and other types of male violence against women can be found: physical violence, sexual violence, psychological violence, economic violence, and stalking. Stalking is a very frequent form of violence, which was recognized as a crime in Italy in 2009 (Art. 612 bis of the Criminal Code, L. n 38 of 2009) and is, together with intimate partner violence, one of the focuses of this research-intervention project. Stalking encompasses a set of persecutory behaviors repeated over time, towards an individual, who develops fear and anxiety, due to the feeling of being constantly followed and/or controlled.

The Zeus Protocol

Signed on April, 5th 2018 by the then Quaestor of Milan, Dr. Cardone, and the President of the CIPM, Dr. Giulini, the Zeus Protocol was created as a tool for first-level secondary prevention and combating gender, domestic and relational violence. It introduces an unprecedented and innovative form of collaboration and cooperation between the Police and the CIPM, within which subjects admonished by the Quaestor for stalking, domestic violence or cyberbullying are invited to follow a treatment path aimed at understanding the social disvalue of their violent actions.

The choice of the name was oriented on the representation of a possessive attitude, with no possibility of containment and reflection on the consequences and risks of one's actions.

Therefore, the Zeus Protocol, which started from the Anticrimine Division of the Milan Police Headquarters as a "pilot" project and as experimentation of an integrated multidisciplinary intervention modality, thanks to the ENABLE project has been extended nationwide, in 10 other Italian provinces, through the signing of similar protocols between Police Headquarters and local CIPMs, also on the impulse of the DAC Circular No. 225/UAG/2019-66981-U of September, 06th 2019.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

It was the so-called "Security Decree" approved by Decree-Law 11/2009 and converted into Law 38/2009, which introduced the criminal offence of "persecutory acts" (Article 612 *bis* of the Italian Criminal Code, commonly referred to as "stalking") and the measure of the warning as an administrative measure *until a complaint is lodged* for the offences 612 *bis* and 612 *ter* of the Italian Criminal Code (Article 8). The meaning of the warning is a formal oral warning by the Quaestor against conduct whose effects are deeply detrimental to the person subjected to it, and an invitation *to behave in accordance with the law*.

According to the article, the reiteration of the conduct by the person, after being admonished, provides for ex officio prosecution, arrest in flagrante delicto, and an increase in sentence.

Subsequently, the instrument of the warning is extended to domestic violence behavior (ex art. 3 Law 119/2013) and cyberbullying (ex art. 7 Law 71/2017), thus recognizing also in these behaviors what we can call "sentinel offence" and in the warning an effective first-level secondary prevention intervention.

Sentinel offences are in fact those behaviors, still characterized by sporadicity, but at risk of escalation from the point of view of the seriousness and harmfulness of the effects they generate in those who suffer them, and found to be sufficiently serious to trigger an administrative sanction; such behaviors need the possibility of being intercepted early, preventively.

A few months after the conclusion of the ENABLE project, the changes in the law contained in the "Caivano Decree" on combating juvenile distress and crime (Decree-Law 123/2023, converted into Law 159/2023) extended the warning to minors over the age of fourteen for offences under articles 581 (battering), 582 (personal injury), 610 (deprivation of liberty), 612 (threats) and 635 (damaging) of the Criminal Code; and for minors between twelve and fourteen years of age (who cannot be charged under the Italian criminal system) for offences provided for by law as crimes punishable by imprisonment of not less than a maximum of five years.

This choice testifies to the recognition of the warning as a tool for the timely and effective interception and intaking of risk situations.

The warning procedure provides that the offended person may report the offending conduct, without initiating criminal proceedings, to the Police Office, which, after assessing the report through an investigative phase, will decide whether or not there are grounds for a warning. The law provides that, for cases of domestic violence or cyberbullying, anyone aware of the facts may request a warning, resulting as a protected source within the administrative procedure.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

The Zeus Protocol enhances and organizes the innovative provision introduced by Art. 3, p.5, Law 119/2013 in the warning procedure: the contextual sending for treatment through a formal invitation to attend at least one interview with an operator at the CIPM, in order to start a preliminary assessment and monitoring process. Although the admonished person is not obliged to attend such an interview, non- participation is reported to the police, so that evaluation can be made regarding the implementation of more incisive prevention measures by the Quaestor. In fact, a parallel can be drawn between the violation of precautionary measures in criminal proceedings, which is a risk factor for the reiteration of deviant conduct, and the refusal to participate in the treatment referral, assessed as a sign or "symptom" of a reduced awareness and accountability of what has happened and what is postponed in the formal act of warning, but also of criticality in identifying and implementing appropriate, functional, adaptive behavior.

The aim of the treatment intervention of the Zeus Protocol is to provide the admonished person with a space for reflection and elaboration with a direct impact on his/her own behavior to the benefit of the offended party, of society and of the person himself/herself.

The Zeus Protocol is in fact a measure that can be defined as "Restorative Justice", because it promotes tools and sensitivity in the community that can promote practices designed to prevent and combat violence, operating outside the penal system but integrating its possibilities of intervention and effectiveness.

Another aspect borrowed from Restorative Justice is the centrality of the protection of the offended person and the reparation of the damage caused, made possible, as in all treatment interventions, by the work of accountability and awareness of the offender.

In the implementation modalities of the Zeus Protocol there is the possibility of contact and meeting with the offended person for a preliminary work of listening, support and assessment of the situation, of the context and of the risks of re-victimization, even secondary. When the offended parts are notified of the issuance of the warning, they can freely express their consent to be contacted by the CIPM, a qualified interlocutor for a monitoring intervention but also, if necessary and requested, to be sent to other specialist victim support services, to psychological support or to a talking group.

The multidisciplinary evaluation, re-evaluation and monitoring work on the treated situations, always within the implementation modalities of the Zeus Protocol, provides for quarterly meetings between the CIPM team and the Police operators, for a comparison on the progress of the

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

treatment paths and an evaluation, case by case, of possible risk or critical factors. The sharing of the outcome and progress of the treatment has nothing to do with its contents but with the way in which the offender participates and adheres to it.

The overall meaning of the treatment is in fact to encourage the admonished person to reflect on his or her own fragility and the consequences of his or her actions with a view to reparation and prevention.

The methodology used to conduct the interviews with the admonished person is inspired by the clinical-criminological interview as a diagnostic and prognostic tool, to get to know the person's functioning, to deepen the motivations, the genesis and dynamics of the admonished conduct, to identify any risk and protection factors with respect to the reiteration or escalation of harmful behavior.

There are two main specificities of the intervention setting: the absence of demand and therefore the treatment *injunction* and the fact that it does not operate within the framework of the criminal prosecution that identifies the subject as a perpetrator; the admonished person is therefore a free citizen who has acted dysfunctional and risky behavior.

The first contact has a mainly informative purpose, aimed at clarifying the objective and functioning of the warning and the Zeus Protocol, defining the setting and the constraints governing it, explaining the function of the operator and treatment.

The preliminary interviews, not defined or definable in number a priori, are space and opportunity to fostering motivation through a mode of encounter and confrontation that holds together welcoming and *questioning*. Welcoming the subject's fragilities, the narration of the self, of the relationship and of the conduct enacted within it, the defenses, minimizations, distortions, even denials; welcoming them in order to be able to question them, through a joint *reconstruction* that gives a new and different meaning to what has been enacted, and that takes into account the reading of the subject but also the contribution of the operator as representative of the community and of the offended party, and of their demands.

At the end of the preliminary interviews, the operator can assess what intervention to propose and what restitution to give the offender. Some people already have a good critical capacity with respect to the conduct they have acted, they understand its effects on the offended party and the social disvalue; these are cases in which the effect of the Police intervention and a short-medium term monitoring work appear adequate and sufficient. Other persons show criticalities in the acquisition of an

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

adequate awareness and accountability, but the time of the preliminaries made it possible to create a motivation and an alliance to work such as to support a short individual pathway aimed at elaborating and consolidating what had emerged, or a referral to an external therapy pathway, supported by ah help demand.

Finally, in some cases, the assessment of the risk of escalation or recidivism or other factors lead the operator to formulate the need to intake of the offender and to include him/her in a medium-long term treatment programme; this is carried out in a group setting, on a weekly frequency, and is designed as a tool for *benevolent control*, the promotion of a critical and reflective attitude, the development of empathy and awareness of the effects of one's behavior, and accountability.

In cases where relevant psychopathological aspects emerge, referral to a specialist service is possible, as a parallel or alternative intervention to treatment.

Statistical data and mapping

Before presenting the results of the research, it is important to introduce some Italian statistical data on warnings. The data come from the Ministry of the Interior (Department of Public Security - Central Directorate of Criminal Police) and were processed by the Central Anticrime Directorate of the State Police.

In 2022, 3,654 individuals were admonished in Italy (up from 2,886 in 2021), of whom 78 percent were found to be of Italian nationality. Furthermore, in 2022 recidivism was 9.7 percent (compared to 14.8 percent in 2021). Regarding the warnings issued in 2022, 703 are reported within Questuras adhering to the Zeus Protocol, while 2,951 are related to Questuras not adhering to the Zeus Protocol. Of the cases related to Questuras adhering to the Zeus Protocol, the recidivism rate is 7 percent, while in Questuras not adopting the Zeus Protocol, the incidence of recidivism is 10 percent.

At the national level, interesting data are found regarding the provincial distribution of the number of warnings. Catania turns out to be the province with the highest number of warnings in 2022 (346), followed by Trento (200), Milan (181), Turin (153), Rome (142), Naples (129), Venice (124), Padua (100), Genoa (82) and Palermo (77). Of the top 10 provinces by number of warnings in 2022, two took part in the ENABLE Project (Milan, Rome and Genova). The ENABLE Project also covered the provinces of Reggio Emilia (49), Cagliari (26), Vicenza (42), Verona (42), Bari (52), Piacenza (9), Savona (57) and Prato (22). The province with the

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

lowest number of warnings in Italy in 2022 is Campobasso (2), followed by Sondrio (3) L'Aquila (3).

In the 12 months prior to the signing of the ZEUS Protocol the recidivism index by province was: Bari (21%), Cagliari (23%), Genoa (10%), Milan (14%), Piacenza (20%), Prato (24%), Reggio Emilia (18%), Rome (12%), Savona (23%), Verona (23%) and Vicenza (24%). After 12 months of signing the ZEUS protocol, the recidivism index by province was: Bari (4%), Cagliari (38%), Genoa (23%), Milan (8%), Piacenza (0%), Prato (24%), Reggio Emilia (16%), Rome (7%), Savona (6%), Verona (20%) and Vicenza (13%).

Objectives

In the context of the scientific activities of the ENABLE Project, the overall objective was to collect information regarding the psychological, social, and value dimensions associated with gender-based violence, both from law enforcement and from instantiating and admonished parties. The information collected represents both a useful national database for benchmarking future similar projects and a measure of the effectiveness of the intervention itself.

More specifically, an initial objective of the research was to collect data from law enforcement agencies regarding variables such as: the knowledge of the instrument of warning, methods of referral to treatment and implementation of multiagency intake, and knowledge of gender-specific forms of victimization (in particular, in relation to stalking and domestic violence). A specific anonymous quantitative survey was developed for this objective.

A second specific objective, aimed at the instant parties, was to investigate some clinical and individual adjustment variables (in terms of depression and posttraumatic symptoms, coping strategies, and risk of victimization). For this objective, a specific anonymous quantitative survey was developed and sent to the instant parties participating in the ENABLE Project.

Finally, a final specific objective was aimed at the admonished, with the purpose of investigating some clinical and individual adjustment variables (in terms of Post-Traumatic Stress Disorder symptoms, coping strategies, risk of recidivism, and adverse childhood experiences). Also for this objective, a specific anonymous quantitative survey was developed and sent to the admonished participating in the ENABLE Project.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Methodology

Participants

A total of 148 officers participated in the Law Enforcement Survey, of whom 56.8% (n = 84) were female, having an average age of 48.5 years (SD = 8.6; range: 26-62) and coming in 45.9 (n = 68) from northern Italy, 22.3% (n = 33) from central Italy, and 29.7% (n = 44) from southern Italy and islands.

Fourteen female individuals (mean age: 41.3 years; range: 19-59years) and 31 male individuals (mean age: 42.8 years; range: 24-56 years) participated in the Instant Part Survey and 31 male individuals (mean age: 42.8 years; range: 24-56 years) in the Admonished Survey. Participants were recruited through Project partners. Anonymity and confidentiality was guaranteed to all participants.

Tools

The following tools were used to construct the Survey addressed to law enforcement agencies:

- the Victim Credibility *Scale* (Page, 2008), to assess the victim's credibility. This instrument was translated into Italian by the research team, using the translation and backtranslation method;
- the *Quantitative Survey with qualitative components to examine law enforcement perspectives and attitudes about* Domestic Violence (Toon & Hart, 2005), to investigate representations and beliefs about domestic violence. This instrument was translated into Italian by the research team, using the translation and backtranslation method;
- the *Secondary Traumatic Stress Scale* (STSS; Bride et al., 2004), to assess the psychological effects of direct exposure to situations connoted by traumatic stress (such as criminal victimization). This instrument was translated into Italian by the research team, using the translation and backtranslation method.

The following tools were used to construct the Survey addressed to the instant parties and the Survey addressed to the admonished:

- Impact of Event Scale - Revised (IES-R; Weiss & Marmar, 1997; Weiss, 2007), for the assessment of Post Traumatic Stress Disorder (PTSD) symptoms. This instrument was used in its Italian adaptation (Giannantonio, 2003);

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

- Beck Depression Inventory - II (BDI-II; Montano & Flebus, 2006; Becket et al., 1961), for the assessment of depressive symptoms. This instrument was used in its Italian validation;
- Coping Orientation to the Problems Experiences-new Italian version (COPE-NVI; Sica, et al., 2008), for the identification of coping strategies. This instrument was used in its Italian validation;
- Juvenile Victimization Questionnaire (JVQ; Finkelhor et al., 2011), for the assessment of victimization in youth. This instrument was translated into Italian by the research team, using the translation and backtranslation method;
- Bright Sky (Vodafone Foundation, Cadmi and State Police, 2020), Italian questionnaire for assessing victimization experiences.

Procedure

All Surveys were constructed using the online platform "Qualtrics" and were distributed to territorial project partners via links. Subsequently, the partners distributed the Surveys to participants, both with regard to Law Enforcement (through SCA) and to admonished and instant parties (through the regional offices of CIPM and ARES). Subsequently, through the Qualtrics platform, the results were collected and analyzed. In all cases, the utmost confidentiality and protection of privacy was ensured, as no personal or sensitive data was collected.

Analysis

Descriptive and correlational analyses were conducted for the Survey targeting law enforcement officers. A comparison analysis of the mean scores obtained (Student's *t-test* for independent samples) between the experimental group (ENABLE training participants) and the control group (non ENABLE training participants) was conducted.

As for the Surveys addressed to the admonished and the instant parties, the results were analyzed by percentage calculation. The results were then compared with each other.

Results

Law Enforcement.

The credibility of the victim

Results on victim credibility in domestic violence cases do not appear to be related to the respondent's age, gender, or length of service. Instead,

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

there are some significant associations between the number of domestic violence and/or stalking cases (followed in the past six months) and the credibility of someone you know ($r = .165$ $p = .050$), a *transgender person* ($r = .254$ $p = .003$), a *teenager* ($r = .260$ $p = .002$), of an older person ($r = .197$ $p = .021$), of a person in prostitution ($r = .245$ $p = .004$), of a man ($r = .294$ $p = .001$), of a professional woman ($r = .270$ $p = .001$), of a married woman accusing her husband ($r = .266$ $p = .002$). This means that the more frequent the respondents' current experience "in the field," and more credibility is given to the victim, regardless of type. Furthermore, for *transgender* victims, there is a further significant correlation between the number of years of experience with respect to domestic violence and/or stalking cases and the credibility accorded to the victim ($r = .185$ $p = .029$). This means that, for *transgender* victims, credibility is also related to years of service experience.

Similarly, results on victim credibility in stalking cases do not appear to be related to respondent's age, gender, or length of service. There are some significant associations between the number of domestic violence and/or stalking cases (followed in the last six months) and the credibility of someone you know ($r = .171$ $p = .046$), a transgender person ($r = .256$ $p = .003$), a teenager ($r = .284$ $p = .001$), of an elderly person ($r = .219$ $p = .010$), of a person in prostitution ($r = .275$ $p = .001$), of a man ($r = .269$ $p = .002$), of a professional woman ($r = .251$ $p = .003$), of a married woman accusing her husband ($r = .273$ $p = .001$). Similar to domestic violence cases, this means that the more frequent the respondents' current "on-the-ground" experience, the more credible the victim appears to be.

Representations and beliefs about domestic violence

Regarding representations and beliefs related to domestic violence and related psychological dynamics, most respondents (64.2%; $n = 95$) believe that domestic violence is a significant problem for the community in which they serve (*versus* 10.2%; $n = 15$). Almost all respondents (94.5%; $n = 140$) do not believe that domestic violence should be considered a private matter rather than a matter for law enforcement (*versus* 0.7%; $n = 1$).

Most respondents (77.6 percent; $n = 115$) do not believe that victims of domestic violence are often as responsible for the events as the person arrested (*versus* 2.1 percent; $n = 3$).

The majority of respondents (51.3%; $n = 76$) do not believe that victims of domestic violence often exaggerate about the extent of violence

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

experienced (*versus* 8.1%; *n* = 12). 39.9% of respondents (*n* = 59) do not take a definite position (neither agree nor disagree).

Most respondents (47.9%; *n* = 71) believed that substance abuse by abusers was a major cause of domestic violence (*versus* 17.5%; *n* = 26).

For the majority of respondents (66.2%; *n* = 98), most domestic violence situations occur because of the abusers' anger management problems (*versus* 12.1%; *n* = 18).

According to the largest number of respondents (75.6 percent; *n* = 112), most domestic violence situations stem from the abusers' need for control and exercise of power over victims (*versus* 4.1 percent; *n* = 6).

The majority of respondents (63.5 percent: *n* = 94) say that mental disorders contribute significantly to domestic violence situations (*versus* 11.5 percent; *n* = 17).

Most respondents (68.8 percent; *n* = 102) believed that one of the main problems of domestic violence was attributable to the many cases of recidivism (*versus* 4.8 percent; *n* = 7).

Regarding representations and beliefs referring to intervention methodology and practice (with particular reference to arresting abusers), 38.6 percent of respondents (*n* = 57) do not believe they need more freedom to decide how to intervene in domestic violence situations (*versus* 16.8 percent; *n* = 25).

The majority of respondents (60.8%; *n* = 90) believe that most victims of domestic violence are responsive and receptive to law enforcement interventions (*versus* 6.1%; *n* = 9).

38.8 percent of respondents (*n* = 59) felt they needed more precise guidelines from superiors on how to behave in domestic violence calls (*versus* 27 percent; *n* = 40), and in addition, 94.4 percent of respondents (*n* = 119) felt more training could help on how to behave in domestic violence situations (*versus* 4.7 percent; *n* = 7).

44% of respondents (*n* = 65) do not believe that, in domestic violence cases, law enforcement should only make arrests when there is clear evidence of injury (*versus* 24.3%; *n* = 36).

38.5% of respondents (*n* = 57) believe it is more likely to make an arrest when there are children in domestic violence cases, who witness the violence (*versus* 27%; *n* = 40).

Most respondents (65.5%, *n* = 97) believe that domestic violence perpetrators should be arrested even when the victim does not want to (*versus* 10.8%, *n* = 16).

The *Secondary Traumatic Stress Scale* (STSS; Bride et al., 2004) was administered to assess the psychological effects of direct exposure to situations connoted by traumatic stress.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

A positive trend (r = .165 p = .055) emerges between intrusive symptoms and the number of domestic violence and/or stalking cases followed in the past six months, meaning that the greater the number of cases followed, the greater the presence of intrusive traumatic symptoms. A higher burden of traumatic stress (and associated symptoms), on the one hand, turns out to be related to the awareness that domestic violence represents a significant problem for the community in which one serves, marked by high levels of recidivism and a certain causal complexity, which imposes strategies of contrast and intervention that go beyond normal practices (resulting in a sense of unpreparedness and difficulty in case management). On the other hand, the development of traumatic symptomatology seems to be associated with evasive attempts to downplay/minimize the real psychological (and social) scope of this phenomenon, considering that many reports of domestic violence instead concern mere verbal family quarrels, which are often traceable to substance abuse problems on the part of the abusers and which go on to occupy too much time and resources for law enforcement officers.

In the context of the ENABLE Project, there was a specific training activity directed to law enforcement officers, oriented on the general issues of gender-based violence and stalking and personal attitudes about these issues. The majority of respondents (85.1%, n = 126) report that they did not participate in ENABLE training.

To better understand whether ENABLE training appears to be associated with a different profile of representations/beliefs between participants *versus* non-participants, analyses were conducted to compare the mean scores obtained (Student's t test for independent samples) between the two groups.

With respect to victim credibility (in cases of domestic violence and stalking) and traumatic psychological effects from exposure to criminal victimization, no significant differences were found between experimental group (the training participants) and control group (the non-training participants). Therefore, the two groups are shown to be similar with respect to the representations that guide the construction of a victim's credibility (of domestic violence and/or stalking) and the traumatic psychological effects (in terms of total score, intrusive symptoms, avoidance symptoms and nervous activation) that may result from exposure to criminal victimization.

Significant differences emerge, however, with respect to representations and beliefs about domestic violence. The following table shows the main descriptive statistics (mean and standard deviation) of the experimental

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

group (the training participants) and the control group (the non-training participants)

Table 1 - *Descriptive statistics (mean and standard deviation) between participants versus non-participants in ENABLE training.*

	ENABLE training participants		Non-Participants in ENABLE Training			
	M	DS	M	DS	t	p
Domestic violence should be considered as a private matter rather than a matter for law enforcement agencies	4,80	0,410	4,58	0,612	-2,022	.050
Victims of domestic violence are often just as responsible for the events as the person arrested	4,40	0,598	4,05	0,822	-2,306	.056
Most victims are safer as soon as they break an abusive relationship	2,45	0,686	2,81	0,890	2,096	.044
Most domestic violence situations stem from the abusers' need for control and exercise of power over victims	2,40	0,681	2,02	0,788	-2,016	.046
One of the main problems of domestic violence is that there are many cases of recidivism	2,65	0,933	2,13	0,765	-2,739	.019
I am more likely to be physically injured by intervening in a domestic violence call than by intervening in a stranger violence call	3,55	0,605	3,24	0,817	-2,017	.052
In domestic violence calls, it is often best to arrest both parties	4,70	0,571	4,32	0,771	-2,598	.014
If the victim of domestic violence shows cooperation, I am more likely to apprehend the abuser	2,90	0,718	3,33	0,969	2,358	.025
More training could help me on how to deal with domestic violence situations	2,65	1,226	1,94	0,727	-2,535	.019

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Regarding representations and beliefs related to domestic violence and related psychological dynamics, our results show that ENABLE training participants, compared to non-participants, report more disagreement with the following statements:

- domestic violence should be regarded as a private matter, rather than a matter for law enforcement (M = 4.80 *versus* M = 4.58; t = -2.022; *gdl* = 142; p = .050);
- victims of domestic violence are often as responsible for the events as the person arrested (M = 4.40 *versus* M = 4.05; t = -2.306; *gdl* = 143; p = .056);
- most domestic violence situations stem from the batterers' need for control and exercise of power over victims (M = 2.40 *versus* M = 2.02; t = -2.016; *gdl* = 143; p = .046);
- one of the main problems of domestic violence is the many cases of recidivism (M = 2.65 *versus* M = 1.94; t = -2.739; *gdl* = 142; p = .019).

While, compared to non-participants, ENABLE training participants report greater agreement that most victims are safer as soon as they break an abusive relationship (M = 2.45 *versus* M = 2.81; t = 2.096; *gdl* = 141; p = .044).

Finally, regarding the belief that most victims are safer as soon as they break off an abusive relationship, only 5 percent of ENABLE training participants strongly disagree/disagree (*versus* 26.8 percent of non ENABLE training participants).

On the other hand, with regard to representations and beliefs referring to the methodology and practice of intervention (with particular reference to arresting abusers) in cases of domestic violence, our results show that training participants, compared to non-participants, report more disagreement with the following statements:

- Be more likely to be physically injured by intervening in a domestic violence call than in a stranger violence call (M = 3.55 *versus* M = 3.24; t = -2.017; *gdl* = 143; p = .052);
- In domestic violence calls it is often better to arrest both parties (M = 4.70 *versus* M = 4.32; t = -2.598; *gdl* = 142; p = .014).

Whereas, compared with non-participants, training participants report greater agreement that if the domestic violence victim shows cooperation, the abuser is more likely to be arrested (M = 2.90 *versus* M = 3.33; t = 2.358; *gdl* = 142; p = .025).

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Instant parties and admonished

A high risk of PTSD was found in 32% of responding admonished (*n* = 31) (*versus* 57% of responding instant parties; *n* = 14). Thus, instant parties have a higher possibility of risk for developing Post Traumatic Stress Disorder than the admonished.

68% of the responding admonished (*n* = 31) scored low in the risk of developing depression (*versus* 54% of the responding instant parties; *n* = 14). In the instant parties, it is possible to find 33% moderate risk and 13% high risk of depression (*versus* 26% and 6% in the admonished). Thus, the instant parties have a higher depression risk than the admonished.

The most frequently used coping strategy by respondent admonishments (n = 31) is possessing a positive attitude (63%), followed by avoidance strategies 19%. In the responding instant parties (n = 14) the most used coping strategy is again positive attitude (34%), which deviates only slightly from the second most used coping strategy, which is problem orientation (33%). In the admonished the coping strategy of problem orientation is the least used (3%), while in the instant parties the least used appears to be avoidance strategies (7%).

Youth victimization

Regarding the risk of juvenile victimization, through the JVQ instrument, it was possible to find that 48% of the responding admonished (*n* = 31) were victims of property assaults at a young age (*versus* 50% of the responding parties; *n* = 14).

3% of the admonished respondents (n = 31) were victims of sexual assault at a young age, while as for the instant parties (n = 14), 36% say they were victims .

Sixty-two percent of admonished respondents (n = 31) have been victims of physical assault at a young age, while as for petitioning parties, 71% of respondents (n = 14) claim to have been victims.

Seventy-seven percent of the responding admonished (*n* = 31) say they were victims of at least one type of maltreatment at a young age (*versus* 57 percent of the responding petitioners; *n* = 14).

Twenty percent of admonished respondents (*n* = 31) claim to have witnessed intimate partner violence at a young age (*versus* 43% of responding admonished parties; *n* = 14).

Forty percent of admonished respondents (*n* = 31) say they witnessed at least one type of violence at a young age (*versus* 43% of responding admonished parties; *n* = 14).

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Forty-seven percent of admonished respondents (*n* = 31) say they were victims at a young age of indirect exposure to violence (*versus* 50 percent of responding petitioners; *n* = 14).

64% of responding instant parties (*n* = 14) say they are afraid or feel anxiety in the presence of their partner; 57% of responding instant parties say their partner's use of physical violence against themselves, their children, or a pet; 57% of responding instant parties consider their partner too jealous and/or controlling; 43% of responding instant parties say their partner has threatened to kill them, and by the same percentage say their partner has used or threatened to use a weapon against them.

14% of the petitioners feel that the situation is getting worse; 7% of the petitioners say they are experiencing economic violence from their partner; while 0% of the petitioners say they have been forced by their partner to perform sexual acts against their will (versus 29% who answered "don't know" and 64% who answered "no").

Discussion and conclusions

With regard to the effect of the ENABLE training, from the significant differences that emerged between the experimental group (the training participants) and the control group (the non-training participants), it is possible to conclude that the training enabled the participants to gain a greater awareness regarding the complexity of the phenomenon of domestic violence, the real understanding of which goes far beyond the context of a "legal fact" to which a sanction corresponds. For ENABLE training participants, it becomes clearer how domestic violence should be considered as the outcome of an articulated and multifaceted psychological dynamic (Kimber et al., 2018), which has deep origins within individuals and couples (Bonura, 2018), and which requires a broader external and social "taking charge". In addition, training participants appear more aware in overcoming some "clichés" referring to domestic violence ("victims are often just as responsible for the acts as perpetrators" or "domestic violence stems from the abusers' need for control and exercise of power"), which make the phenomenon simple, downplaying its magnitude. In this sense, training participants show that they have accurately internalized a reading of the dynamics of gender-based violence as an exercise of power within intimate relationships, in concordance with the Duluth Model's "wheel of power and control" metaphor (Pence & Paymar, 1993).

Increased awareness and understanding of the psychological characteristics of domestic violence also make ENABLE training

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

participants more open to countering and restorative interventions: in fact, through targeted and specialized actions, domestic violence can be effectively curtailed and taken care of, interrupting or at least limiting the phenomenon of recidivism.

These representational aspects then spill over into practical operations, methodology and intervention practice. Again, training participants are shown to be more competent in "reading" the criminological dynamics inherent in domestic violence situations, in which arresting abusers cannot be considered the sole and definitive to the problem. In fact, the request turns out to be more complex and requires a careful assessment of the individuals involved.

Finally, ENABLE training participants say they are more confident in dealing with domestic violence situations and feel less of a need for further pathways.

Following the administration of the surveys for admonished and instant parties, one can, first of all, find limited and uneven numbers between the two groups. The results regarding the risk of developing post-traumatic stress disorder are very interesting, with the instant parties being more exposed to this possibility than the admonished. These results agree with the literature; in fact, there are several studies that have shown a significant association between the development of PTSD and being a victim of Intimate partner violence (Lily, 2010; Mignone et al., 2017), and stalking (Dardis et al., 2017). Specifically, the study by Dardis et al. found that women who were victims of stalking by their partners were 2.5 times more likely to develop PTSD than women who had never experienced stalking by their partners.

From the results, it can also be seen that the instant parties are more likely to develop a depressive disorder, compared to the admonished. This result agrees with the literature from which it is possible to find that intimate partner violence significantly increases, in victims, the risk of developing mood disorders, such as depression and suicidal ideation (Okuda et al., 2011; Afifi et al., 2009).

Regarding coping strategies, it is interesting to note heterogeneity in response modes, but also how in both groups positive attitude was reported as the most used strategy, and how the second most used response mode in the instant parties is also the least used in the admonished (problem orientation) and the second most used in the admonished is also the least used in the instant parties (avoidance strategies). In contrast, the literature shows that there is a tendency for victims to implement the avoidance coping strategy and notes that coping strategies related to emotional

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

regulation are those that most moderate intimate partner violence and the development of PTSD (Mignone et al., 2017; Lilly, 2010),

With the administration of the JVQ instrument related to victimization in youth, it is interesting to note that both instant parties and admonished were significantly victimized by physical assault and other unspecified types of mistreatment. The results diverge significantly in the variable related to sexual assault, which show greater victimization of this crime among the instant parties. This finding is congruent with the hypothesis of intergenerational transmission of violence, and in particular with the numerous data in the literature indicating that being a victim of assault (especially in the family context) exposes one to a subsequent increased risk of enacting violent acts in turn (cf. Widom & Wilson, 2014).

Limits and future directions

The most obvious limitations were found as a result of the administration of the surveys related to admonished and instant parts. As mentioned in the introduction, quantitative surveys were not planned within the project, but semi-structured, qualitative focus groups were planned. Given the considerable difficulty in recruiting participants, two quantitative surveys were added so as to go to increase the data pool. Although the surveys were distributed online, under the premise of ensuring total anonymity and protection of the participant, a relatively eximious number of data were retrieved.

Another limitation noted was the numerical difference of instant parties and admonished, evidenced by a significant gap that made the two groups uneven (N instant parties = 14; N admonished = 31).

A third limitation concerns the use – in some cases – of instruments for which no official Italian validation and adaptation is available. This is because the variables measured by these scales are of an extremely specialized nature, and no other instruments were available. The translation and back-translation procedure met all the criteria required by good scientific practice.

As reported, the territorial services showed considerable difficulty in recruiting participants, especially with regard to the requesting parties who either were not present within the service or did not agree to participate in the research. This difficulty is definitely to be taken into consideration for further research on this topic, so that useful strategies can be found in such a way as to go and increase the amount of data.

The study by Dardis (2017), cited earlier, found that women who experienced intimate partner violence and stalking by their ex-partner

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

were 4.2 times more likely than women victims of intimate partner violence without stalking to develop Post Traumatic Stress Disorder. One could try to understand whether there is, in light also of the results obtained in this project, a correlation between the development of Post Traumatic Stress Disorder and the relationship the victim has with their stalker (partner, ex-partner or stranger).

In the future, research on this topic should even more effectively take up the issue of delving into the history of early unfavorable experiences in people at risk of committing domestic violence or stalking, delving into the issue of "intergenerational transmission of violence" and identifying the specific risk profiles that would allow for early interception of people prone to enacting gender-based violence.

Funding Statement

The research-intervention referred to in this article is funded by the EU REC-RDAP-GBV-AG-2020 "ENABLE" fund.

References

Afifi, T. O., MacMillan, H., Cox, B. J., Asmundson, G. G., Stein, M. B. & Sareen, J. (2009). Mental health correlates of intimate partner violence in marital relationships in a nationally representative sample of males and females. *Journal of Interpersonal Violence,* 24, 1398-1417. DOI: 10.1177/0886260508322192.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. & Erbaugh, M. D. (1961). An Inventory for Measuring Depression. *Archives of General Psychiatry*. *4*(6), 561-571. DOI:10.1001/archpsyc.1961.01710120031004.

Bonura, M. L. (2018). *What kind of violence: knowing and coping with violence against women*. Trento: Erickson study center editions.

Bride, E. B. Robinson, M. M., Yegidis, B. L & Figley, C. (2004). Development and Validation of the Secondary Traumatic Stress Scale. *Research on Social Work Practice, 14*(1):27-35. DOI: 10.1177/1049731503254106.

Courtois, C. A. (1988). *Healing the incest wound: Adult survivors in therapy.* New York: Norton.

Dardis, C. M., Amoroso, T., & Iverson, K. M. (2017). Intimate partner stalking: Contributions to PTSD symptomatology among a national sample of women veterans. *Psychological Trauma: Theory, Research, Practice, and Policy*, 9(1), pp. 67-73. DOI: 10.1037/tra0000171.

Danieli, Y. (1988). Confronting the unimaginable: Psychotherapists' reactions to victims of the Nazi Holocaust. In J. P. Wilson, Z. Harel, &B. Kahana (Eds.),

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Human adaptation to extreme stress: From the Holocaust to Vietnam, (219-238). New York: Plenum.

Dutton, M. A., & Rubinstein, F. L. (1995).Working with people with PTSD: Research implications. In C. R. Figley (Ed.), *Compassion fatigue: Coping with secondary traumatic stress disorder in those who treat the traumatized* (pp. 82-100). NewYork: Brunner/Mazel.

Figley, C. R. (Ed.). (1995). *Compassion fatigue: Coping with secondary traumatic stress disorder in those who treat the traumatized*. New York: Brunner/Mazel.

Finkelhor, D., Hamby, S. L., Ormrod, R. & Turner, H. (2005). The Juvenile Victimization Questionnaire: reliability, validity, and national norms. *Child Abuse Neglect, 29*(4), 383-412. DOI: 10.1016/j.chiabu.2004.11.001.

Garbarino, F., & Giulini, P. (2022). The treatment of offenders: the multidisciplinary approach in combating gender-based violence. *Child Maltreatment and Abuse*, 24(3), 37-60.

Garbarino, F., & Giulini, P. (2020) Violence in close relationships. CIPM's restorative justice experience. In B.Pezzini & A. Lorenzetti, (Eds.) *Gender violence from the Rocco Code to the Code Red. An itinerary of plural reflection through the complexity of the phenomenon*. Turin: Giappichelli.

Giannantonio M. (2003). *Psychotraumatology and emergency psychology.* Salerno: Ecomind.

Haley, S. A. (1974). When the patient reports atrocities: Specific treatment considerations of the Vietnam veteran. *Archives of General Psychiatry*, 30, 191-196.

Herman, J. L. (1992). *Trauma and recovery*. NewYork: Basic Books.

Toon, R. & Hart, B. (2005). *Layers of Meaning: Domestic Violence and Law Enforcement Attitudesin.* Arizona, Department of Justice, Office on Violence Against Women: Washington, DC.

McCann, I. L., & Pearlmann, L. A. (1990). Vicarious traumatization: A framework for understanding the psychological effects of working with victims. *Journal of Traumatic Stress*, 3, pp. 131-149.

Mignone, T, Pagani, E. Mahadeo, M., Klostermann, K. & Jones R. A. (2017). PTSD and Intimate Partner Violence: Clinical Considerations and Treatment Options. *J Addict Med Ther Sci, 3*(1), 001-006. DOI: 10.17352/2455-3484.000018.

Montano, A., & Flebus, G. B. (2006). Presentation of the Beck Depression Inventory-Second edition (BDI-II): Confirmation of the bifactorial structure in an Italian population sample. *Cognitive and Behavioral Psychotherapy, 12*(1), pp. 67-82.

Lilly, M. M. (2010). Intimate Partner Violence and PTSD: The Moderating Role of Emotion-Focused Coping. *Violence and Victims, 25*(5), 604-610. DOI: 10.1891/0886-6708.25.5.604.

Okuda, M., Olfson, M., Hasin, D., Grant, B.F., Lin, K.H., & Blanco, C. (2011). Mental health of victims of intimate partner violence: Results from the

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

National Epidemiologic Survey on Alcohol and Related Conditions. *Psychiatric Services*, 62, 959-962. DOI: 10.1176/ps.62.8.pss6208_0959.
Page, A. D. (2008). Judging women and defining crime: police officers' attitudes toward women and rape. *Sociological Spectrum, 28*(4), 389-411, DOI: 10.1080/02732170802053621.
Pence, E., & Paymar, M. (1993). *Education groups for men who batter: The Duluth model*. Springer Publishing Company.
Sica, C., Magni, C., Ghisi, M., Altoè, G., Sighinolfi, C., Chiri, L. R., & Franceschini, S. (2008). Coping Orientation to the Problems Experiences-new Italian version (COPE-NVI). *Cognitive and Behavioral Psychotherapy*, *14*(1), 27-53.
Kimber, M., Adham, S., Gill, S., McTavish, J., & MacMillan, H. L. (2018). The association between child exposure to intimate partner violence (IPV) and perpetration of IPV in adulthood-A systematic review. *Child abuse & neglect,* 76, 273-286.
Weiss, D. S. (2007). The Impact of Event Scale: Revised. In J. P. Wilson & C. S.-k. Tang (Eds.), *Cross-cultural assessment of psychological trauma and PTSD*, (pp. 219-238). Springer Science + Business Media. DOI: 10.1007/978-0-387-70990-1_10.
Weiss, D. S., & Marmar, C. R. (1997). The Impact of Event Scale-Revised. In J. P. Wilson & T. M. Keane (Eds.), *Assessing Psychological Trauma and PTSD* (pp. 399-411). New York: Guilford.
Widom, C. S., & Wilson, H. W. (2014). *Intergenerational transmission of violence. Violence and mental health: Its manifold faces*, (pp. 27-45).

Sitography

ARES Association APS (https://www.centroares.com/).
Italian Center for the Promotion of Mediation (www.cipm.it).
CRIdee (https://centridiricerca.unicatt.it/CRIdee).
State Police - Anti-Crime Division (https://www.interno.gov.it/it/ministero/dipartimenti/dipartimento-pubblica-sicurezza/direzione-centrale-anticrimine-polizia-stato).
https://www.cipm.it/enable/ (project landing page).

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Cracking the Code: Examining Psychometric Rigor of the Provider Decision Process Assessment Instrument (PDPAI) among Residents' Trainees and Expert Physicians

Sana Younas*, Saeeda Khanum**

Department of Behavioral Sciences,
School of Social Sciences and Humanities,
National University of Sciences and Technology, Islamabad Pakistan.
*e-mail: sanayounasmphil@gmail.com.
**e-mail: saeeda.khanum@s3h.nust.edu.pk.

Submitted: 29.03.2024 - **Accepted:** 29.07.2024

Available online: 28.11.2024

Abstract

The realm of healthcare decision-making remains inadequately explored, specifically in assessing the psychometric characteristics of tools like the Provider Decision Process Assessment Instrument (PDPAI). This study aims to fill this void by examining decisional conflict among resident trainees and experienced consultant physicians. We approached a total of 347 physicians using a convenient sampling method from tertiary care hospitals. The analysis encompassed (i) factorial validity of PDPAI through confirmatory factor analysis (CFA) and evaluating the single group CFA models and (ii) multigroup CFA models, (iii) examining factorial invariance among residents' trainees and experienced physicians' groups, (iv) Rasch analysis assessing the individual item impact on the subdomains, (v) internal consistency (vi) convergent and discriminant validity. The bi-factor model adequately fit the data as all factor loadings (0.44-0.70) were statistically significant ($p < 0.05$). The bifactor model supported the global construct or the sub-domains as suitable measurement models. The PDPAI showed invariance for use across two physician groups. Physicians encountered greatest difficulty in item "I was clear what treatment would be best for this patient." [MNSQ Infit/Outfit: 1.327/1.278] and found the easiest item "It was easy to identify all of the considerations that affect the decision" [0.902/0.869]. Adequate internal consistency was revealed through

Sana Younas & Saeeda Khanum/ *Ricerche di Psicologia, 2024, Vol. 47*
ISSNe 1972-5620, Doi: 10.3280/rip2024oa18309
Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Cronbach and Omega coefficient values. Convergent and discriminant validity of PDPAI was supported by correlating with team decision making questionnaire and compassion fatigue respectively. The PDPAI's validated cross-group invariance highlights it's applicability to a diverse range of physician groups, guiding tailored interventions.

Keywords: healthcare physicians, factorial validity, Rasch analysis, Internal consistency, convergent and discriminant validity

Introduction

This study endeavors to fill a crucial void in existing literature by scrutinizing the psychometric attributes of the 12-item Provider Decision Process Assessment Instrument (PDPAI) (Dolan, 1999) within the cohorts of resident trainees and experienced consultant physicians. The identified critical gap pertains to the insufficient understanding of decisional conflict among these groups – conflict stemming from uncertainty surrounding potential action plans, as articulated by Han et al., (2019). Despite the existence of a number of decision evaluation scales (for e.g., PDPAI (Dolan, 1999), Decision Attitude Scale (Barry et al., 1997), the Satisfaction with decision making process questionnaire (Sainfort & Booske, 2000) etc.), a thorough examination into their robustness, psychometric properties, and application in various health settings and populations remains conspicuously absent. There is a lack of research that requires an investigation, into verifying decision assessment scales especially in healthcare settings to ensure their accuracy and consistency. This gap is highlighted by the number of studies focusing on the characteristics of PDPAI, among both early career and experienced physicians, also known as consultants or attending physicians. These professionals offer guidance, supervision, and specialized knowledge to trainees during their training period (Younas et al., 2023: Younas & Khanum, 2024a, b). By exploring the intricacies of decision making in these circles, we may gain insights into reducing decision related conflicts and ultimately improving patient outcomes in healthcare settings.

Decisional conflict results when a person is uncertain about a potential course of action (Moure et al., 2023; O'Connor et al., 2002). Additionally, decisional conflict is aggravated when people compromise on values in order to choose a course of action or when they anticipate feeling re-morseful for not selecting certain options (Liu et al., 2023). The main

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

behavioral manifestations of decisional conflict are expressed uncertainty about possibilities, expression of unfavorable consequences of alternatives, and a propensity for delaying decisions (Dhami & Mandel, 2022). Decisional conflict is contributed by factors such as lack of support or external pressure, unclear personal values, and limited knowledge. Decisional conflict can be effectively reduced by decision supporting interventions. People feel more informed when they are given information about options, and side effects (Liao et al., 2023). Detailed explanations of outcomes, including their effects on the body, mind, and spirit, are a key component of value-clarification strategies. People are also urged to evaluate the significance of these outcomes for themselves. Guidance during the stages of collaborative decision-making promotes support in decision-making. As a result, uncertainty resulting from these adaptive aspects decreases, giving the impression that the option was better (Légaré et al., 2006; O'Donnell et al., 2023). This improved choice is distinguished by a sense of better knowledge, agreement with personal beliefs, increased likelihood of adherence, and higher levels of satisfaction. The beneficial effects of decision-supporting treatments on decisional conflict and the associated changeable dimensions are strongly supported by empirical data (O'Connor, 1995; Wendler & Rid, 2011).

By addressing the presumptive causes of this conflict, decision aids have the potential to reduce decisional conflict (O'Conner, 2010). For instance, a thorough examination of the options and a thorough investigation of the prospective outcomes can help to offset the lack of information. Physicians'confidence in the informed nature, alignment with personal values, and feasibility of their decisions is gauged through the efficacy of their decision making. The existing scale evaluates how comfortable physicians feel about their decisions. The importance of examining medical decision making processes in healthcare environments is emphasized due to the potential influence on patient results and healthcare procedures.

The Health Physician's Version of the Decisional Conflict Scale (DCS)

Dolan (1999) first adapted the DCS among 22 healthcare workers (Trees et al., 2017). Numerous studies emphasized the relevance of decisional conflict as a marker of high-quality decision-making among medical professionals. Accordingly, the adapted version of the DCS, called the Provider Decision Process Assessment Instrument (PDPAI), was developed to measure decisional conflict among medical professionals. According to Zimmer-Watson et al. (2008), PDPAI fills a

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

significant research gap, by gauging individual medical professionals' perceptions of the decision process. The adapted instrument comprised four more items, in addition to the original 16, resulting in a 20-item instrument where eight of the 20 items were the same as in DCS (i.e., items 9, 10, 13, 18, 19, 21, 22, and 25). It showed an acceptable construct validity and reliability, as revealed through the Cronbach's alpha of 0.90. Another study (Honmg et al., 2003) translated the PDPAI into French and administered it to 34 family physicians in Quebec, resulting in a Cronbach's alpha coefficient of 0.82. Furthermore, nine out of the original 16 items of DCS were used in an Australian study that investigated the ability of family physicians to help patients decide about prostate-specific antigen screening (as cited in Pecanac et al., 2018). This study also showed consistency of the instrument as reveled through the Cronbach's alpha value ($\alpha = 0.81$).

However, there are some shortcomings of the scale. First, Dolan (1999) recruited only 112 participants resulting in limited generalizability and implications of the sample. Second, the author did not explore the factor structure of the scale to establish the construct validity rather used Spearman correlation and found negative correlation between decision conflict and satisfaction. Third, Dolan's study has been cited in 67 research articles, however, to the best of our knowledge, none of the studies worked on factor structure of the scale. Some of the review studies, however, suggested that the PDPAI instrument exhibits good internal consistency in terms of psychometric quality, but also suggested that validity has not been sufficiently examined (Scholl et al., 2011; Simon, Loh, & Härter, 2007). Hence, it was necessary to validate the scale in an Asian culture before using it.

Cultural traditions in collectivistic nations such as Pakistan, emphasize group decision-making that frequently affect medical decisions. For instance, decisions about how to manage elderly patients may involve extensive family influence, influencing both end-of-life and medical care options (García & Garasic, 2021). The PDPAI has been validated to ensure that it appropriately captures the subtleties of decision difficulty encountered by physicians while respecting the cultural setting in which they work. Physicians may encounter potential challenges influencing their ability to make well-informed decisions such as a lack of access to literature or information gaps. The validation of this scale assures that the evaluation properly captures the ethical issues inherent in physicians'decision-making processes.

The objective of this investigation was to evaluate the psychometric characteristics of the 12-item PDPAI through the application of both classical test theory and item-response theory. We examined the

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

congruence of the domains using confirmatory factor analysis for (i) single factor PDPAI (ii) two factor PDPAI and (iii) bi-factor PDPAI. We also examined the PDPAI for measure equivalency in early career and experienced physicians. Since, it has been emphasized that years of experience in uncertain medical situations helps one to make better decisions (Falzer, 2018; Klein, 2015). Individuals having more experience use more intuitive thinking, cues, and recognize the situation through pattern formation as compared to the young physicians who have less experience of dealing with uncertain situations in medical settings (Epstein, 2011; Ruzsa et al., 2020). Young trainee physicians having less experience may get more stressed in uncertain and time pressured situations, thus having more conflict with their decisions as compared to experienced physicians (for instance, consultants) who are less stressed and are more satisfied with their decisions. Additionally, Rasch modeling was used alongside confirmatory factor analysis. Wilson et al. (2006) noted, Rasch modeling provides a direct assessment of the relationship between respondents' positions and the placement of items on the latent variable scale. Contrarily, confirmatory factor analysis, as outlined by Strauss and Smith (2009), examines the connections between components. Omega reliability, convergent and discriminant validity of the PDPAI were also assessed.

Methods

Sample

Ethical approval was obtained prior to data collection from ethical review board of the University. The research was a descriptive cross-sectional survey. Data collection started in March 2022 and ended till January 2023 from all major tertiary care hospitals in the Potohar region located northeast of Pakistan. Convenient sampling technique was used to approach the participants. Greater sample size is considered better for validation of a measure. Hence, the 10:1 minimum standard (10 cases per parameter) was used to draw the sample size (Hair et al., 2010). Total 347 practitioners participated in the study. Among them, (n = 180) were residents' trainees who were referred as early career physicians while (n = 167) were senior and experienced physicians including classified consultants (Assistant Professors, Associate Professors, and Professors). Their age ranged from 24-66 years (M = 32.066; SD = 7.513). Resident trainees had experience between 1 and 4 years while senior physicians have 6 and more years of experience.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Measures

Provider Decision Process Assessment Instrument (PDPAI) developed by Dolan (1999) to measure the healthcare provider assessment of decision-making. The instrument is in English, hence did not require translation as Pakistani physicians can understand and speak English well. The assessment comprises a total of 12 items, with scoring conducted on a five-point Likert scale. It explores the perceived complexity present in situations involving decision-making and acknowledges the intricate interactions between variables that may make medical decisions more challenging.

Additionally, Dolan's scale evaluates clinicians'knowledge of patients'preferences and values during the decision-making process, further emphasizes the significance of patient-centered care. It examines how much medical practitioners value shared decision-making, respects patients'autonomy, and involves patients in treatment conversations. Some of the examples of items include item 3: "*I fully understand the patient's views regarding the important issues in making this decision";* item 11: *I am satisfied with the decision that was made*; item 12: *I am satisfied with the process used to make the decision was as good as it could be*.

Response options are rated on a scale of 1 to 5 where 5 shows "strongly agree" and 1 shows "strongly disagree". Items 6, 5, 4, 2, and 1 are negatively scored hence, they are reversed before calculating the total score of items 1-12. The maximum possible score range lies between 12 to 60. The scale has a sufficient alpha value ($\alpha = 0.87$). Participants were instructed via the demographic information sheet to recall a recent critical case they had managed. They were then asked to respond to the scale items based on their perception of that specific medical case.

Two additional measures were selected that were expected to be associated with decision-making; a team-decision-making questionnaire (Batorowicz & Shepherd, 2008) that measures the overall team support of healthcare providers in decision-making. This scale consists of 12 items and utilizes a seven-point Likert scale for rating. The response option for score 1 is "never" while the response option for score 7 is "to a large extent". The internal consistency of the measure is >.90 with all positively worded items. The scale is expected to correlate positively with PDPAI showing evidence of convergent validity. An additional sub-scale of professional quality of life i.e., compassion fatigue (CF) developed by Stamm and adapted by Galiana et al. (2020) was chosen to establish the discriminant validity of PDPAI. Compassion fatigue, also known as secondary or vicarious trauma, can be risk for healthcare providers who get

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

exposed to patients' trauma, emergencies, and critical cases daily (Stamm, 2010). Physicians experience compassion fatigue because they deal with patients'trauma, emergencies, and serious cases daily. Chronic stress can result from the emotional toll of witnessing suffering coupled with the duties of the work. Compassion fatigue among healthcare personnel can be caused by a variety of factors, including long work hours, the difficult nature of medical decisions, and the pressure to strike a balance between empathy and professional detachment (Hui et al., 2023). It's a complex issue that goes beyond providing for patients'urgent needs to include the general emotional and mental health of individuals working in the medical field. This subscale employs a five-point Likert scale, with the option "Never" assigned a score of 1 and "Very Often" assigned a score of 5. The scale has a Cronbach alpha value of 0.82. It is expected that decision conflict will correlate positively with compassion fatigue.

Procedure

The study's aims were clearly explained during the recruitment procedure, and each participant gave their written informed consent before being included. Throughout the data collection procedure, confidentiality and privacy were prioritized. The necessary institutional review boards were consulted for ethical approval, which underlines the dedication to respecting ethical norms in research. They were given thorough information about the research including possible risks and benefits, and were given the assurance that participating was optional. To protect the participants'rights and wellbeing, all aspects of the process were conducted by ethical standards.

Participants were informed about the study's purpose, procedures, risks, and benefits. Written informed consent was obtained from all participants, ensuring their understanding and voluntary participation. Anonymity was ensured by using unique codes to protect participants'identities. Data was only accessible to the research team and was stored on password-protected computer. Additionally, to prevent individual identity, data was presented in aggregate form. Participants were told that data would be retained for a period of five years and then will be disposed of permanently deleting electronic files.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Data Analysis

SPSS version 26 was used to calculate the values of descriptive statistics. Group differences were assessed through independent sample T-test among the two physician groups. Multicollinearity diagnostics were assessed by calculating the Variance Inflation Factor (VIF) that showed values less than 5 indicating that multicollinearity was not an issue with the study variables. Moreover, we conducted the confirmatory factor analysis (CFA) through AMOS version 22 statistical software for three models (i) single-factor PDPAI (ii) two-factor PDPAI and (iii) bi- factor PDPAI to assess the factor structure of decision conflict. CFA is used to validate any existing measure because it helps in verification of the measure's factor structure with accuracy. Additionally, CFA is supportive in analysis of any underlying relationship between observed and latent factors. Fit indices were obtained with no residual dependency (Padgett & Morgan, 2021).

Factorial invariance of the PDPAI was examined across two groups of physicians' senior physicians (consultants/experts) (n = 167) and residents' trainees'/early career physicians (n = 180). Once the optimal factor solution was identified for both groups, the least restrictive configural invariance model was employed. By carefully investigating the distribution of fixed and free model parameters, this model evaluated the equivalence of the overall factor structure between the two groups without imposing any equality restrictions (Widaman & Olivera-Aguilar, 2023). To guarantee equality between groups, the metric invariance model subsequently set restrictions on each item's factor loading. The goal of this research was to ascertain whether the correlations between variables and factors were the same in both the early career residents in training and experienced physician groups. The aim was to investigate if every PDPAI item is consistently loaded into the same factor in both groups. Metric invariance, also known as weak measurement invariance, signifies consistent measurement units on the scale, indicating a shared understanding of the items among individuals in both groups. By limiting item intercepts, the scalar invariance model investigates if items share identical intercepts (item means) across both groups. Scalar invariance, commonly known as robust measurement invariance indicates that item scores in both groups adhere to a consistent measurement metric and identical scalar, enabling comparisons of factor means between the groups. The absence of scalar invariance implies the potential presence of systematic bias in response patterns between the two groups (Leroux et al., 2023). The stringent level of constraint is represented by the factor variance invariance model, commonly referred to as structural invariance.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

To determine if the relationships between latent components are consistent across two groups of physicians, it sets extra limits on factor variances and covariances (Sass & Schmitt, 2013). This degree of invariance examines if the PDPAI scores for the two groups fall within the same range and whether the connection between the components is constant across physicians'groups.

Goodness of fit for the bi-factor model structure through structural invariance, factor variance invariance, metric invariance, and configural invariance were examined using these fit indices (i) Root Mean Square error of Approximation (RMSEA) (i) Comparative Fit Index (CFI) (iii) Standardized Root Mean Square Residual (SRMR). The subsequent cut-off values were employed as indicators of a well-fitting model; (1) RMSEA < .08 (2) CFI > 0.95 (3) SRMR < .08 (Hu & Bentler, 1999).

The present study also applied item response theory through Rasch modelling (Wright & Stone, 1999) on Jamovi analysis software program. An analysis of the outfit mean square (OUTFIT MNSQ) and infit mean square (INFIT MNSQ) statistics was undertaken (Smith, 2001). The analyses examine and assess the degree to which the observed data and the Rasch model's expected values correspond (Smith, 2000). This makes it easier to assess how much each component defines a shared construct. As suggested by Wright (1994), we deemed items to be "fit" if their MNSQ is between 0.6 and 1.4. Items with fit statistics between 0.5 and 1.5, however, may still be regarded as useful for measurement. A low score denotes simpler observations or less diversity in response patterns, while a high score denotes difficult observations or a wide range of responses to an item.

Additionally, discriminant validity was determined through Average Variance Extracted (AVE) and Maximum Shared Variance (MSV) criteria (Voorhees et al., 2016). The average variance of a factor should be higher than the average variance that it shares with all other factors (Farrell et al., 2009). To put this another way, the maximum shared variance (MSV) should be lower than the average variance (AVE) extracted. Also, the convergent validity was established through Pearson Product Moment correlation. PDPAI was correlated with team decision- making scale to assess convergent validity whereas the discriminant validity was assessed by correlating PDPAI with compassion fatigue. Since it is expected that PDPAI scale would positively correlate with team decision-making scale however lack of strong correlation is expected between PDPAI and compassion fatigue. The PDPAI's observed distinctiveness in the correlation pattern with its target construct, along with a modest link

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

with compassion fatigue, are consistent with the findings of Fernández-Miranda et al. (2023) reinforcing the discriminant validity of the scale.

Finally, the Cronbach alpha was calculated for the overall sample and subscales as well as across two physician groups to determine the internal consistency reliability using SPSS software. In addition, McDonald Omega estimation (Ω) was calculated using JASP software freely available from the University of Amsterdam. Also, split-half reliability was calculated using the Spearman brown formula.

Results

Descriptives

Descriptive statistics are calculated for mean (*M*), standard deviation (*SD*), skewness (skew), and kurtosis (ku) for the scale. The values for conflict domain are (*M* = 16.06, *SD* = 4.10, skew = 0.13; ku = -0.11) while the values for satisfaction domain are (*M* = 25.75, *SD* = 4.03; skew= 0.06; ku = -0.07). Skewness and kurtosis lie within the acceptable range of ±1 showing the data is normally distributed.

Tab. 1 - *Group Differences on PDPAI scale and subscales*

Variables	Residents Trainee physicians		Consultant Physicians		t	p	95% CI		Cohen's d
	M	SD	M	SD			LL	UL	
PDPAI	41.09	5.99	43.22	6.41	-3.14	.00	-3.45	-8.08	0.34
D_Conflict	14.98	4.19	15.86	4.97	-1.73	.08	-1.35	.11	0.19
D_Satisfaction	26.10	4.17	27.35	4.09	-2.78	.00	-2.14	-3.68	0.30

Note. PDPAI = Provider Decision process Assessment Instrument, D_conflict = Decision conflict, D_Satisfaction = Decision satisfaction

Consultant physicians scored high on overall PDPAI scale as well as on decision satisfaction domain. Furthermore, non-significant differences are observed on decision conflict domain among the two groups (Table 1) Cohen's d is also too small and has almost negligible effect size. However, a modest effect size is indicated by a Cohen's d of 0.34 and 0.30 in case of group differences across overall scale and satisfaction domain respectively. This indicates that although there is a discernible difference

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

between the two groups, it is not very significant. While the effect sizes are statistically significant, it is important to assess their practical significance. In practical terms, depending on the situation, even slight variations can have significant implications. For instance, small adjustments made to decision-making procedures may eventually result in better patient outcomes in a healthcare context. If there are not many differences between the groups, there may not be much of an impact on practice or policy. Stakeholders ought to contemplate whether the discernible disparities result in feasible modifications or enhancements that warrant the necessary exertion of time or resources. Future studies could explore these differences with larger sample sizes or different methodologies to verify whether the modest effect sizes persist or if more pronounced differences emerge. Researchers must also consider any additional variables or factors such as social and environmental variables that have an impact on effect sizes. Investigating these additional variables can offer a more thorough comprehension of the noted variations and their consequences.

Single Group CFA models

The measurement indices revealed a poor fit for the unidimensional decision conflict domain for either group of physicians. The values of fit indices were far from the acceptable range (AGFI = .35; GFI = .46; RMSEA = .27; IFI = .47; TLI = .35; CFI = .47). The two-factor structure of PDPAI comprising of "satisfaction" and "conflict" factors, demonstrated satisfactory fit in both early career physicians (CFI = .91, RMSEA = .08, SRMR = .05) and experienced physicians (CFI = .92, RMSEA = .04, SRMR = .06) groups. While this model improved on the unidimensional model, it still did not fully encapsulate the complexity of the PDPAI. Finally, the bifactor model, encompassing a singular general factor and two specific domain factors (conflict and satisfaction), demonstrated a good fit for both groups (CFI = .95, RMSEA = .05, SRMR = .03). As the bi-factor model was the only one exhibiting a satisfactory fit for both physician groups, examinations of structural and measurement invariance were solely performed on the bi-factor model. Significant standardized factor loadings were observed in the overall sample for the bi-factor model, ranging from 0.51 to 0.60 for the "conflict" factor and 0.44 to 0.70 for the "satisfaction" factor (Table 2).

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Tab. 2 - *Factor loadings from Trainee Residents and Specialists baseline models of the PDPA*I

PDPAI-items	Postgraduate-Trainee Residents (*n* = 180)		Consultants (Experts) (*n* = 167)	
	Unstandardized	Standardized	Unstandardized	Standardized
Conflict				
1.	1.00	.58	1.00	.60
2.	1.08*	.54	1.02*	.55
4.	1.06*	.51	1.03*	.56
5.	0.76*	.51	0.78*	.60
6.	0.74*	.55	0.75*	.49
Satisfaction				
3.	1.00	.44	1.03	.50
7.	1.43*	.56	1.11*	.55
8.	1.20*	.58	0.89*	.60
9.	1.69*	.53	1.55*	.60
10.	0.99	.58	1.23	.66
11.	1.11	.67	1.00	.64
12.	1.00	.63	0.98	.70

Note: To establish the measurement scale for the underlying latent variable, the factorial loading of the first item was constrained to a value of 1,
*$p < .05$

Multigroup CFA models

We examined the fit indices of the factor variance invariance, scalar, metric, and configural, models across decision making of experienced physicians (consultants) and early career physicians (residents' trainees) for the bi-factor structure of the PDPAI which is presented in Table 3. For invariance testing, a change of less than or equal to 0.01 indicates that the invariance holds in ΔCFI. A change of less than or equal to 0.015 is

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

considered acceptable for ΔRMSEA. For metric invariance, a change of less than or equal to 0.03 is acceptable for ΔSRMR; for scalar invariance, a change of less than or equal to 0.01 is acceptable. Additionally, for model fit (CFI $\geq$ 0.95, RMSEA $\leq$ 0.06, SRMR $\leq$ 0.08, TLI $\geq$ 0.95) indicates that the same factor structure is valid across groups. For metric invariance, overall model fit is reassessed and in comparison, to configural model using ΔCFI ($\leq$ 0.01) and ΔRMSEA ($\leq$ 0.015) is considered acceptable. Minimal changes in fit indices suggest that factor loadings are equivalent across groups. For scalar invariance, overall model fit is again reassessed compared to metric model using ΔCFI ($\leq$ 0.01) and ΔRMSEA ($\leq$ 0.015). For factor variance invariance, overall model fit is reassessed; compared to scalar model using ΔCFI ($\leq$ 0.01) and ΔRMSEA ($\leq$ 0.015). Minimal changes in fit indices suggest that factor variances are equivalent across groups.

Configural invariance

Without imposing any equality constraints on the two factor structure of model, configural invariance was investigated for the experienced and early career physicians groups. For both groups, the values of CFI, IFI, and SRMR were in acceptable range showing that the baseline two-factor model adequately fits the data. Statistically significant unstandardized factor loadings are obtained for both groups that are also in the same direction (λs = 0.74–1.69, *ps* < 0.001). This provides additional support of the bi-factor structural model across configural invariance.

Metric invariance

To assess metric invariance, all factor loadings were constrained to be equivalent across the two groups. The findings indicate that factor loadings remained consistent across the two groups, supported by a well-fitted model for the data (IFI > 0.90, CFI > 0.90, SRMR and RMSEA < 0.08) hence, metric invariance was achieved. The next step was to compare the less restrictive configural invariance model to the metric invariance. Consequently, results revealed descriptive (all Δ values <0.01) or non-statistical (p > .05) differences in model fit. Hence, the weak measurement invariance was satisfied and metric invariance model was considered a more fitting match for the data.

Scalar invariance

We examined the scalar invariance by constraining item intercepts and factor loadings to equivalence across the two groups. Both the item intercepts and factor loadings were invariant across the two groups as model fit was obtained for the scalar invariance (IFI > 0.90, CFI > 0.90, SRMR and RMSEA < 0.08). While making a comparison of the

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

constrained model to the less restrictive metric invariance model, non-statistical descriptive differences (all Δ values < 0.01) were observed in model fit. Hence, there was a significant measurement invariance.

Factor variance invariance

We investigated the factor variance invariance by applying equivalence constraints across the two groups on item intercepts, factor loadings, and factor variance invariance. Acceptable model fit was obtained as revealed through the values of IFI, CFI and RMSEA. While comparing it to the less restrictive model of scalar invariance, a decline in model fit was observed (ΔCFI = -.016, ΔRMSEA = -.002, ΔSRMR = -.008). This indicates a dearth of factor variance invariance. Hence, the model with scalar invariance was deemed to be more optimal fit for the data.

Tab. 3 - *Goodness-of-fit indices for models testing scalar, configural, metric, and factor variance invariance of the PDPAI*

Model	Df	P	CFI	RMSEA	SRMR	Δdf	Δ p	Δ CFI	Δ RMSEA	Δ SRMR
Configural	104	< .001	.930	.073	.071	---	---	---	---	---
Metric	105	< .001	.920	.062	.075	1	.979	-.001	-.011	.004
Scalar	118	< .001	.930	.061	.078	13	.395	.01	-.001	.003
Factor Variance Invariance	120	< .001	.924	.059	.070	2	< .001	-.016	-.002	-.008

Rasch Analysis

The findings of the Rasch analysis suggested a generally favorable match (both infit and outfit) between the observed and model expected data (Table 4). The most challenging item found by participants was PDPAI item 3 ("I *was clear what treatment would be best for this patient*"), while item 7 was found to be the easiest one. Furthermore, person reliability statistics revealed 0.824 which is a good fit. In our Rasch analysis, items 1, 2, 3, 4, 5, and 6 were identified as difficult. Items with poor fit have the potential to skew measurements and diminish the instrument's validity. Significantly misfitting items may not be appropriate to all respondents or may measure something different from the intended concept. However, A DIF analysis was performed to check whether these items displayed variations in difficulty or discrimination between subgroups. This analysis helped us ensure that the items are functioning fairly across different groups. No significant DIF was detected in these difficult items (1, 2, 4, 5, 6), indicating that they may not be equally challenging or discriminative for the two subgroups. If the items showed significant DIF, this could introduce bias, making it difficult to

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

compare scores across the two groups. However, there are several reasons for misfit of the items. (i) Items could measure something different or less relevant than expected if they don't align well with the intended concept. The item content may not be clear, or the language may be confusing, causing this mismatch. (ii) Secondly, if respondents perceive the items differently, poor fit may result in inconsistent results. These biases may result from response styles that differ, cultural variations, or differing levels of familiarity with the item content. (iii) Thirdly, Misfit can occur when items with extremely high or low difficulty levels do not align well with the overall distribution of respondent skills. For example, an item may not fit well into the Rasch model if it is too simple or difficult in comparison to the other items. (iv) Lastly, the fit of some items may be impacted by their redundancy or excessive complexity. Items that are very complex or comparable to one another may not offer any new or distinctive insights into the construct being measured.

Furthermore, the Wright map is a very useful sort of statistic which has been outputted in Figure 1. On the right-hand side are the items which have been sorted based on their difficulty. For example, item no. 9, 10, 7, 8, 12, and 11 cluster together and are the easiest ones which fall at the bottom then there is a little bit of gap between these items and the rest of the items. On the left-hand side is the distribution of people which has been demonstrated using bar graphs. Hence, there is a good distribution in items as well as people. Since there are people from all levels of ability and the ability ranges from around -2 logits to the estimated ability of larger than 4. Here, it is also essential to mention that as per the person reliability statistic which is 0.824; indicates that we have been able to estimate the ability levels of our test takers with 82.4 percent of precision.

Tab. 4 - *MNSQ Infit and Outfit values indicating the conformity of the observed data to the model's expected data*

Items no.	MNSQ INFIT	MNSQ OUTFIT
Dec_PA1_1	0.995	0.980
Dec_PA2_1	1.087	1.099
Dec_PA3_1	1.327	1.278
Dec_PA4_1	1.118	1.103
Dec_PA5_1	1.099	1.106
Dec_PA6_1	1.129	1.133
Dec_PA7_1	0.902	0.869
Dec_PA8_1	0.891	0.872

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Dec_PA9_1	0.941	0.925
Dec_PA10_1	0.892	0.922
Dec_PA11_1	0.937	0.942
Dec_PA12_1	1.121	1.100

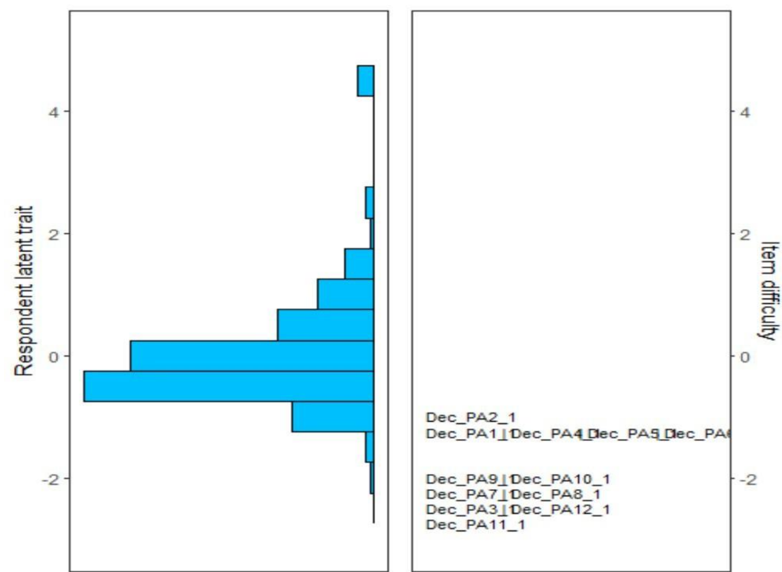

Fig. 1 - *A Graphical Representation of Item Response Theory through Wright Map*

Internal Consistency and McDonald's Omega Reliability

Internal consistency for the PDPAI scale was adequate for the total sample (α = .79), early career physicians' group (α = .87), and experienced physicians' group (α = .78). The factor 1 consisting of conflict items revealed alpha (α = 0.86) while Factor 2 consisting of satisfaction items revealed alpha (α = 0.81) for early career physicians (residents) while Factor 1 revealed alpha (α = 0.81) and factor 2 revealed alpha (α = 0.81) for experienced (consultants) physicians**.** The overall McDonald's omega reliability was found to be (Ω = 0.76), however for satisfaction items it was (Ω = 0.85) and for conflict items, the omega coefficient was (Ω = 0.81). Additionally, using the Spearman-Brown formula, which adjusts for the fact that we only used half of the scale at a time, we estimated the reliability of the full scale. The split-half reliability was found to be 0.77, which is satisfactory. This reliability indicates that the two halves of the scale are very consistent with each other.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Convergent and Discriminant validity

The value of AVE is found above the cut-off value of 0.50 (Henseler et al., 2016). The factor loadings of the items for each respective construct helped in determining the convergent validity. The percentage of AVE for both the factors; satisfaction (AVE = 0.511, MSV = 0.373) and decision conflict (AVE = 0.542, MSV = 0.440) was in acceptable range. The construct's individual correlation with each of the other factors was smaller than the square root of average variance extracted AVE values for each sub-scale. Considering this, it is suggested that the AVE of each individual factor was more than the MSV.

Moreover, we also established criterion-related validity by Pearson correlation. Decision satisfaction was correlated with team support in decision-making ($r = .484^{**}$, $p < .00$) showing a positive correlation. The effect size is moderate. In practical terms, this suggests that change in decision satisfaction moderately predict change in team support. Additionally, decision conflict correlated positively with compassion fatigue ($r = .382^{**}$, $p < .00$) with moderate effect size. Hence, evidence of convergent validity is supported. Additionally, decision satisfaction correlated negatively with compassion fatigue ($r = -.099^{*}$, $p < .05$) showing evidence for discriminant validity. However, the strength of the effect size is almost negligible in size.

Discussion

A significant lacuna in the psychometric assessment of the decision assessment scale was identified by the literature, which paid scant attention to important factors including measurement invariance, convergent validity, and discriminant validity. To bridge this gap, the current study carried out a thorough measurement invariance test as well as a rigorous psychometric assessment using both single and multigroup Confirmatory Factor Analysis (CFA). Additionally, the investigation of convergent and discriminant validity deepens our comprehension of the robustness of the PDPAI and provides a nuanced contribution to the existing body of literature.

The PDPAI demonstrated a favorable fit within a sample comprising resident trainees and (experienced) consultant physicians, establishing itself as a reliable and valid measure for this particular demographic. Acceptance of a bi-factor model, alongside well-fitted factor variance invariance, scalar, metric, and configural equivalent models, underscores the appropriateness of the instrument for the dataset. The bifactor model affirms the suitability of employing either the subdomains or a global measurement model of the PDPAI.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Factorial invariance analysis on the two groups of resident trainees (early career) and consultants (experienced) physicians indicated that metric and scalar invariance showed a better fit than that of the configural or factor variance invariance models, suggesting that PDPAI can be applied to both experienced physicians and early career trainees. The factor variance invariance of the PDPAI was not achieved, and it was possible to distinguish the correlational associations between the latent variables of the two groups. The instrument remains valid for comparing the average levels of the latent construct between groups due to achieved scalar invariance. However, we are unable to compare the degree of variability or dispersion of the construct between groups due to the inability to demonstrate factor variance invariance. Comparing the latent construct's distribution or degree of variability between groups becomes difficult in the absence of factor variance invariance. Since the instrument does not measure variability consistently across groups, differences in construct variability cannot be evaluated with confidence. Future studies may consider looking at possible causes of group-to-group variability in the construct. This can entail determining if particular items contribute disproportionately to the variance or whether response pattern variations between groups have an impact on the dispersion of the concept. Furthermore, future researchers may look into different measurement models or methods that can provide a better alignment of factor variation between groups. This can entail improving the instrument's fairness and usefulness for a variety of respondent groups.

Although the structural invariance was not present, we reemphasize that the scores within the PDPAI were appropriate for different levels of physicians (early career residents and more experienced consultants). Measurement invariance (both metric and scalar) is thought to be essential for cross-group comparisons. Cheung (2004) describes invariance as 'measuring the same construct' as similar scores between the two groups are less likely to be a result of sample bias, and more likely to represent true group differences.

Rasch analysis confirmed a good fit between empirical and expected data and are in line with convincing evidence from current research (e.g., Paceco-Colón et al., 2019). This validates the importance of accuracy in measuring instruments to capture the variability that exists in decision-making processes. The problems mentioned, such as, respondents' inability to define what would be best for patients (PDPAI question 3), complexity of their decision process (PDPAI question 2) and easiness to find aspects to influence the decision (PDPAI item 7), were previously found in the relevant literature. Zhou and Xu (2023) found that medical decision-making is 'an extremely complex process at its core'. Moral

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

distress in relation to clinical ethics and end-of-life decision-making is a significant problem in healthcare practice and a very distressful experience for those involved. The importance of cognitive fluency for decision-making was found by Schwarz and colleagues in 2021. Taking into consideration the aforementioned evidence, and its robustness, it is possible to assume that each of the answered items of the PDPAI revealed even subtle and very small shifts in decision-making processes among physicians. This finding demonstrates one of the PDPAI's main implications to enhance the practical and field research as well as to strengthen its psychometric properties. Another implication refers to potential improvements in interventions created to tackle and alleviate decisional conflict among cancer patients and their families.

Furthermore, the internal consistency of the PDPAI scale revealed substantial alpha coefficients in both groups of the physicians as well as in the total sample. Strong McDonald's omega reliability and consistency in the factor structures for the satisfaction and conflict items reinforce the internal coherence of the scale. These findings demonstrate how well the PDPAI captures the nuanced nature of healthcare professionals' decision-making experiences.

The study found each individual factor's AVE was higher than the MSV, supporting discriminant validity and indicating that each construct is assessing a distinctive and particular component of the phenomenon under study. Thus, the claim that the measurement tool effectively distinguishes between various conceptions is supported by the fact that items inside a construct are more likely to agree than things from other constructs. This study strengthens the measurement model's validity by showing that the constructs are sufficiently distinct and do not overly overlap, adding to the assessment tool's overall robustness. Moreover, criterion-related validity is confirmed by Pearson correlations. The correlation between decision satisfaction and team support highlights the link between successful decision-making and collaborative settings (Wang et al., 2023). The positive relationship between decision conflict and compassion fatigue also emphasizes how decisional difficulties affect caregiver wellbeing (Cocker & Joss, 2016). Positive decision experiences may help to lessen the effects of caregiver fatigue, according to the negative association between decision satisfaction and compassion fatigue. These results confirm the validity of the assessment technique by highlighting the complex interactions between decision-making, team support, and compassion fatigue.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Limitations of the Study

We used convenience sampling which limits the generalizability of the study findings. This is especially true for physicians who practice in different areas, in different types of healthcare facilities, or from different cultural backgrounds. It is critical to consider our results in light of this sampling strategy. We contrasted our results with those from other studies (e.g., Wu et al., 2022) that used similar sampling strategies to lessen this constraint. This comparative study showed that, despite the sampling limitations, our results are consistent with more general trends observed in the field. However, more robust sampling methods, such as random or stratified sampling are recommended for future researchers to enhance the generalizability of the findings. Moreover, follow-up studies using different methodologies (for instance, focus groups or qualitative interviews) could help expand and validate our initial findings.

The study did not explicitly account for cultural factors in the analysis, which is a limitation that needs to be addressed in future research. Cross-cultural studies should be conducted to assess the validity and reliability of the PDPAI. These studies should include diverse samples to capture the cultural variability in decisional conflict. Modification of items as per the cultural relevance is essential which may involve adding culturally specific examples, rephrasing of items, or adjusting response options. Moreover, cultural variables could be included in the analysis to examine how cultural differences impact decisional conflict. This could involve comparing results across different cultural groups and identifying culturally specific patterns.

Conclusion

The PDPAI emerges as a reliable and valid measure for usage among healthcare physicians. The overall measure of decisional conflict, the sub-domains of the scale, and individual items all demonstrated a satisfactory congruence with the collected data. Evaluating decisional conflict stands as a crucial endeavor to gauge the challenges encountered by health physicians in the decision-making process. The PDPAI can be used in future research projects aiming at improving decision-making processes to examine if such improvements help in reduction of decision conflicts among physicians.

Practical Implications

Enhancing the decision-making process by reducing decisional conflict presents a significant problem for physicians. The practicality of totally eliminating uncertainty was criticized by Nelson and colleagues (2007) who said, “for several decisions, the objective of entirely eliminating or even

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

reducing uncertainty is merely impractical" (p. 615). However, it is crucial to understand that within the framework of the decisional conflict, uncertainty just represents one aspect of decisional conflict (O'Conner, 1995). Healthcare professionals can address different subdomains of decisional conflict notwithstanding the inherent ambiguity in difficult medical decisions. They can do this by educating patients, offering assistance, and soliciting their values in an effort to reduce medical decisional conflict.

It is crucial to take into account the effects of decisional conflict in the context of healthcare decision-making procedures given the potential difficulties faced by healthcare professionals who are involved in decision-making. Evaluating decisional conflict as a key indicator is advisable given the complex dynamics and pressures involved in making important decisions for patients. An analogous consideration for what might be referred to as a healthcare professional's decisional dilemma – a situation in which the complexity and importance of healthcare decisions place a heavy load on the decision-makers – exists, building on the idea Decisional conflict is a crucial metric for assessing interventions supporting healthcare professionals in navigating complex healthcare decision-making. Drawing from Netzer and Sullivan (2014), measuring decisional conflict illuminates the challenges faced by professionals, providing a key gauge for intervention effectiveness.

References

Barry, M. J., Cherkin, D. C., YuChiao, C., Fowler, F. J., & Skates, S. (1997). A randomized trial of a multimedia shared decision-making program for men facing a treatment decision for benign prostatic hyperplasia. *Disease Management and Clinical Outcomes*, *1*(1), 5-14.

Batorowicz, B., & Shepherd, T. A. (2008). Measuring the quality of transdisciplinary teams. *Journal of Interprofessional Care*, *22*(6), 612-620. DOI: 10.1080/13561820802303664.

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods*, *22*(3), 507-526. DOI: 10.1037/met000007.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255.

Cocker, F., & Joss, N. (2016). Compassion fatigue among healthcare, emergency and community service workers: A systematic review. *International Journal of Environmental Research and Public Health*, *13*(6), 618.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Dhami, M. K., & Mandel, D. R. (2022). Communicating uncertainty using words and numbers. *Trends in Cognitive Sciences*, *25*(6) 514-526 DOI: 10.1016/j.tics.2022.03.002.

Dolan, J. G. (1999). A method for evaluating health care providers'decision making: the Provider Decision Process Assessment Instrument. *Medical Decision Making*, *19*(1), 38-41.

Epstein, S. (2011). Intuition from the perspective of cognitive-experiential self-theory. In Plessner H., Betsch C., & Betsch T. (Eds.), *Intuition: In judgment and decision making* (pp. 23-37). Hoboken: Taylor & Francis.

Falzer, P. R. (2018). Naturalistic decision making and the practice of health care. *Journal of Cognitive Engineering and Decision Making*, *12*(3), 178-193. DOI: 10.1177/1555343418773915.

Farrell, A. M., & Rudd, J. M. (2009). *Factor analysis and discriminant validity: A brief review of some practical issues*. Anzmac.

Fernández-Miranda, G., Urriago-Rayo, J., Akle, V., Noguera, E., Mejía, N., Amaya, S., & Jimenez-Leal, W. (2023). Compassion and decision fatigue among healthcare workers during COVID-19 pandemic in a Colombian sample. *Plos One*, *18*(3), e0282949.

Galiana, L., Oliver, A., Arena, F., De Simone, G., Tomás, J. M., Vidal-Blanco, G., ... & Sansó, N. (2020). Development and validation of the Short Professional Quality of Life Scale based on versions IV and V of the Professional Quality of Life Scale. *Health and Quality of Life Outcomes*, *18*(1-12). DOI: 10.1186/s12955-020-01618-3.

García, A., & Garasic, M. (2021). Ethical issues concerning informed consent in translational/clinical research and vaccination. In *Cross-Cultural and Religious Critiques of Informed Consent,*(pp. 11-17). Routledge.

Hair, J. F., Anderson, R. E., Babin, B. J., & Black, W. C. (2010). *Multivariate data analysis: A global perspective,* (7). Pearson Upper Saddle River.

Han, P. K., Babrow, A., Hillen, M. A., Gulbrandsen, P., Smets, E. M., & Ofstad, E. H. (2019). Uncertainty in health care: Towards a more systematic program of research. *Patient Education and Counseling*, *102*(10), 1756- 1766.

Hong, S., Malik, M. L., & Lee, M. K. (2003). Testing configural, metric, scalar, and latent mean invariance across genders in sociotropy and autonomy using a non-Western sample. *Educational and Psychological Measurement*, *63*(4), 636-654. DOI: 10.1177/0013164403251332.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: a Multidisciplinary Journal*, *6*(1), 1-55. DOI: 10.1080/10705519909540118.

Hui, L., Garnett, A., Oleynikov, C., & Boamah, S. A. (2023). Compassion fatigue in healthcare providers during the COVID-19 pandemic: a scoping review protocol. *BMJ open*, *13*(5), e069843.

Klein, G. (2015). A naturalistic decision making perspective on studying intuitive decision making. *Journal of Applied Research in Memory and Cognition*, *4*(3), 164-168. DOI: 10.1016/j.jarmac.2015.07.001.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Légaré, F., O'Connor, A. M., Graham, I. D., Wells, G. A., & Tremblay, S. (2006). Impact of the Ottawa Decision Support Framework on the agreement and the difference between patients'and physicians'decisional conflict. *Medical Decision Making*, *26*(4), 373-390.

Leroux, E. J., Rizeq, J., & Skilling, T. A. (2023). Measurement invariance of the Youth Self-Report across youth who have committed sexual and nonsexual offenses. *Psychological Assessment*, *35*(10), 821-829.

Liao, Y. L., Wang, T. J., Su, C. W., Liang, S. Y., Liu, C. Y., & Fan, J. Y. (2023). Efficacy of a decision support intervention on decisional conflict related to hepatocellular cancer treatment: a randomized controlled trial. *Clinical Nursing Research*, *32*(1), 233-243.

Liu, Y., Wang, X., Wang, Z., Zhang, Y., & Jin, J. (2023). Ethical conflict in nursing: A concept analysis. *Journal of Clinical Nursing*, *32*(15-16), 4408-4418. DOI: 10.1111/jocn.16563

Moure, M., Jacobsen, J. B., & Smith-Hall, C. (2023). Uncertainty and climate change adaptation: a systematic review of research approaches and people's decision-making. *Current Climate Change Reports*, *9*(1), 1-26. DOI: 10.1007/s40641-023-00189-x

Nelson, W. L., Han, P. K., Fagerlin, A., Stefanek, M., & Ubel, P. A. (2007). Rethinking the objectives of decision aids: a call for conceptual clarity. *Medical Decision Making*, *27*(5), 609-618. DOI: 10.1177/0272989X07306780.

Netzer, G., & Sullivan, D. R. (2014). Recognizing, naming, and measuring a family intensive care unit syndrome. *Annals of the American Thoracic Society*, *11*(3), 435-441.

O'Connor, A. M. (2010). *User Manual-Decisional Conflict Scale. 1993.* See [https://decisionaid. ohri. ca/eval. html# DecisionalConflictScale].

O'Donnell, D., O'Donoghue, G., Ní Shé, É., O'Shea, M., & Donnelly, S. (2023). Developing competence in interprofessional collaboration within integrated care teams for older people in the Republic of Ireland: A starter kit. *Journal of Interprofessional Care*, *37*(3), 480-490. DOI: 10.1080/13561 820.2022.2075332.

O'Connor, A. M. (1995). Validation of a decisional conflict scale. *Medical Decision Making*, *15*(1), 25-30. DOI: 10.1177/0272989X9501500105.

O'Connor, A. M., Jacobsen, M. J., & Stacey, D. (2002). An evidence-based approach to managing women's decisional conflict. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, *31*(5), 570-581.

Pacheco-Colón, I., Hawes, S. W., Duperrouzel, J. C., Lopez-Quintero, C., & Gonzalez, R. (2019). Decision-making as a latent construct and its measurement invariance in a large sample of adolescent Cannabis users. *Journal of the International Neuropsychological Society*, *25*(7), 661-667. DOI: 10.1017/S1355617719000341.

Padgett, R. N., & Morgan, G. B. (2021). Multilevel CFA with ordered categorical data: A simulation study comparing fit indices across robust estimation methods. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(1), 51-68. DOI: 10.1080/10705511.2020.1759426.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Pecanac, K. E., Brown, R. L., Steingrub, J., Anderson, W., Matthay, M. A., & White, D. B. (2018). A psychometric study of the decisional conflict scale in surrogate decision makers. *Patient Education and Counseling*, *101*(11), 1957-1965. DOI: 10.1016/j.pec.2018.07.006.

Ruzsa, G., Szeverenyi, C., & Varga, K. (2020). Person-and job-specific factors of intuitive decision-making in clinical practice: results of a sample survey among Hungarian physicians and nurses. *Health Psychology and Behavioral Medicine*, *8*(1), 152-184. DOI: 10.1080/21642850.2020.1741372.

Sainfort, F., & Booske, B. C. (2000). Measuring post-decision satisfaction. *Medical Decision Making*, *20*(1), 51-61.

Sass, D. A., & Schmitt, T. A. (2013). Testing measurement and structural invariance: Implications for practice. In *Handbook of quantitative methods for educational research* (pp. 315-345). Brill.

Scholl, I., Koelewijn-van Loon, M., Sepucha, K., Elwyn, G., Légaré, F., Härter, M., & Dirmaier, J. (2011). Measurement of shared decision making-a review of instruments. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, *105*(4), 313-324. DOI: 10.1016/j.zefq.2011.04.012.

Schwarz, N., Jalbert, M., Noah, T., & Zhang, L. (2021). Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. *Consumer Psychology Review*, *4*(1), 4-25.

Simon, D., Loh, A., & Härter, M. (2007). Measuring (shared) decision-making-a review of psychometric instruments. *Zeitschrift für ärztliche Fortbildung und Qualität im Gesundheitswesen-German Journal for Quality in Health Care*, *101*(4), 259-267. DOI: 10.1016/j.zgesun.2007.02.029.

Smith Jr, E. V. (2001). Evidence for the reliability of measures and validity of measure interpretation: a Rasch measurement perspective. *Journal of Applied Measurement*, *2*(3), 281-311.

Smith, R. M. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement*, *1*(2), 199-218.

Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *Annual Review of Clinical Psychology*, *5*, 1-25.

Trees, A. R., Ohs, J. E., & Murray, M. C. (2017). Family communication about end-of-life decisions and the enactment of the decision-maker role. *Behavioral Sciences*, *7*(2), 36.

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science*, *44*, 119-134. DOI: 10.1007/s11747-015-0455-4.

Wang, S., Li, L., Liu, C., Huang, L., Chuang, Y. C., & Jin, Y. (2023). Applying a multi-criteria decision-making approach to identify key satisfaction gaps in hospital nurses' work environment. *Heliyon*, *9*(3)e14721.

Wendler, D., & Rid, A. (2011). Systematic review: the effect on surrogates of making treatment decisions for others. *Annals of Internal Medicine*, *154*(5), 336-346. DOI: 10.7326/0003-4819-154-5-201103010-00008.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Widaman, K. F., & Olivera-Aguilar, M. (2023). Investigating measurement invariance using confirmatory factor analysis. *Handbook of structural equation modeling*, 367-384.
Wilson, M., Allen, D. D., & Li, J. C. (2006). Improving measurement in health education and health behavior research using item response modeling: comparison with the classical test theory approach. *Health Education Research*, *21*(suppl_1), i19-i32.
Wright, B. D. (1994). Reasonable mean-square fit values. *Rasch Meas Transac*, *8*, 370.
Wright, B. D., & Stone, M. H. (1999). *Measurement Essentials*. Wilmington, DE: Wide Range.
Wu, D., Yang, T., Herold, F., Hall, D. L., Mueller, N., Yeung, A., ... & Zou, L. (2022). Validation of the 4-Item and 10-Item uncertainty stress scale in a community-based sample of chinese adults. *Psychology Research and Behavior Management*, 2803-2813.
Younas, S. & Khanum, S. (2024a). Examining the Role of Stress and Team Support in Decision-Making Under Uncertainty and Time Pressure. *MDM Policy and Practice* (in press).
Younas, S. & Khanum, S. (2024b). Unraveling the Predictive Role of Work Rules on Compassion Satisfaction and Career Satisfaction among Professionals of Obstetrics and Gynaecology: The Mediating Effect of Team Support. *Journal of Professions and Organization, 11*(3)1-15. DOI: 10.1093/jpo/joae011.
Younas, S., Khanum, S., & Qamar, A. H. (2023). Decision making among residents in training of obstetrics and gynecology: A qualitative exploration in Pakistani context. *PLoS ONE 18*(11), e0287592. DOI: 10.1371/journal.pone.0287592.
Zettel-Watson, L., Ditto, P. H., Danks, J. H., & Smucker, W. D. (2008). Actual and perceived gender differences in the accuracy of surrogate decisions about life-sustaining medical treatment among older spouses. *Death Studies*, *32*(3), 273-290.
Zettel-Watson, L., Ditto, P. H., Danks, J. H., & Smucker, W. D. (2008). Actual and perceived gender differences in the accuracy of surrogate decisions about life-sustaining medical treatment among older spouses. *Death Studies*, *32*(3), 273-290. DOI: 10.1080/07481180701881230.
Zhou, J., & Xu, X. (2023). The difficulty of medical decision-making: should patients be involved?. *Hepatobiliary Surgery and Nutrition*, *12*(3), 407-409.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Intelligenza emotiva dell'insegnante e relazione insegnante-alunno nella scuola primaria. Uno studio esplorativo

Teacher emotional intelligence and teacher-student relationship in primary school. An exploratory study

Giordana Szpunar*°, Chiara Mannola*, Eleonora Cannoni*, Madalina Grigore*, Anna di Norcia*

*Dipartimento di Psicologia dei Processi di Sviluppo e di Socializzazione, Università degli Studi di Roma "La Sapienza", Piazzale Aldo Moro 5, 00185 Roma.
° E-mail: giordana.szpunar@uniroma1.it.

Ricevuto: 13.05.2024 **- Accettato:** 25.07.2024

Pubblicato online: 28.11.2024

Riassunto

Questo studio ha lo scopo di indagare il ruolo dell'intelligenza emotiva di docenti di scuola primaria, nella qualità della relazione insegnante-studente. La novità di questo lavoro consiste nell'uso di metodologie diversificate per dare voce sia agli insegnanti sia agli alunni allo scopo di indagare il nesso tra intelligenza emotiva dei docenti e la qualità della relazione insegnante-studente. In particolare, ai docenti sono stati somministrati due questionari self report mentre con gli alunni (età media 8,67; DS=.59) è stato utilizzato lo strumento del disegno. Si tratta, in particolare, di un primo studio esplorativo condotto su un piccolo campione composto da 6 insegnanti di classi terze e quarte e da 79 bambini. I risultati, ottenuti dai questionari e dai disegni, mettono in evidenza che le insegnanti con livelli di EI più alti hanno anche migliori relazioni con gli studenti. Invece confrontando il punto di vista di docenti e alunni sulla loro relazione si evidenzia una correlazione moderata.

G. Szpunar et al. / *Ricerche di Psicologia, 2024, Vol. 47*
ISSNe 1972-5620, Doi: 10.3280/rip2024oa18870

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Parole chiave: intelligenza emotiva; scuola primaria; relazione insegnante-studente; disegno.

Abstract

This study explores the role of emotional intelligence in primary school teachers and its impact on the quality of the teacher-student relationship, considering both teachers' and students' perspectives. The novelty of this work is the use of diversified methodologies to give voice to both teachers and students to investigate the link between teachers' emotional intelligence and the quality of the teacher-student relationship. Specifically, two self-report questionnaires were administered to the teachers, while the students (average age 8.67; SD=0.59) used the drawing. This is primarily an exploratory study conducted on a small sample of 6 teachers, from third and fourth-grade classes, and 79 children. The results, drawn from the questionnaires and the drawings, indicate that teachers with higher levels of intelligence also tend to have better relationships with their students. Additionally, a comparison of the perspectives of teachers and students on their relationships reveals significant similarities.

Keywords: emotional intelligence; primary school; teacher-student relationship; drawing.

1. Introduzione

L'ambiente scolastico rappresenta uno dei contesti più importanti per lo sviluppo cognitivo, emotivo e relazionale dell'individuo. Nella scuola primaria l'insegnante è una figura di riferimento e di identificazione estremamente rilevante. La qualità della relazione dell'insegnante con i propri allievi (STR) determina in modo decisivo il benessere e il miglioramento degli esiti di sviluppo di questi ultimi (Roorda et al., 2011). In particolare, un rapporto con l'insegnante caratterizzato da vicinanza, calore, affetto e basso livello di conflitto sembra avere un effetto positivo sullo sviluppo socio-emotivo del bambino (Cefai e Cavioni, 2014), sui suoi progressi nell'apprendimento (Downer, Rimm-Kaufman e Pianta, 2007), sul suo comportamento (Meehan, 2004), sulla relazione con i pari (Pianta, 1999, Molinari e Melotti, 2010). Al contrario, un rapporto con l'insegnante caratterizzato da conflitto e/o dipendenza è correlato negativamente con la regolazione emotiva e il comportamento (Birch e Ladd 1997, 1998; Murray e Greenberg, 2000) e con i risultati scolastici (Hamre & Pianta, 2001).

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Puertas Molero e colleghi (2019) sottolineano l'importanza dell'intelligenza emotiva (EI) nella professione dell'insegnante. Avere una buona percezione dell'EI è un fattore fondamentale per mantenere una salute mentale ottimale, perché le persone con alti livelli di EI hanno un migliore senso di autoefficacia, evitano la sensazione di burnout professionale e permettono un miglioramento delle relazioni con tutti gli attori del contesto educativo. Questo facilita e migliora sia la performance lavorativa dei docenti sia i processi di apprendimento-insegnamento e, quindi, il benessere e la qualità dell'apprendimento degli studenti.

Inoltre, l'intelligenza emotiva influenza lo sviluppo delle abilità sociali ed emotive come riconoscere e gestire le emozioni, prendere decisioni responsabili, sviluppare la cura e la preoccupazione per gli altri, stabilire relazioni positive e affrontare efficacemente i conflitti (Jennings e Greenberg, 2009). Tali abilità possono essere utili nella formazione di relazioni positive tra insegnante e studente (Gunter et al., 2012).

In letteratura sono pochi gli studi che combinano la percezione dell'EI degli insegnanti con le relazioni con i propri studenti. Così come sono poche le ricerche che prendono in considerazione il punto di vista di insegnanti e degli studenti per confrontarli in merito alla loro relazione (Poulou, 2017).

Le ricerche che hanno preso in considerazione la variabile dell'EI degli insegnanti della scuola primaria, infatti, si sono focalizzate soprattutto sul suo rapporto con il burnout (Mancini et al. 2022), l'autoefficacia (Valente et al., 2020), la leadership (Gómez-Leal et al., 2022), il benessere (Puertas Molero et al. 2019).

La novità del presente lavoro, un primo studio esplorativo condotto con un piccolo gruppo di partecipanti, è nell'indagare il nesso tra intelligenza emotiva dei docenti e la qualità della relazione insegnante-studente attraverso l'uso di diversi strumenti.

1.1 L'intelligenza emotiva

L'intelligenza emotiva è un costrutto relativamente giovane e piuttosto complesso e nella letteratura scientifica non se ne può rintracciare una definizione univoca e condivisa (De Caro e D'Amico, 2008).

Mayer e Salovey (1990) sono i primi a tentare una formalizzazione dell'intelligenza emotiva definendola, nel loro primo modello, il *mental ability model*, come l'abilità di comprendere e monitorare i sentimenti propri e altrui e di utilizzarli come informazioni per guidare il pensiero ed il comportamento. Questa definizione esprime perfettamente l'idea secondo cui l'emozione renda il pensiero più intelligente, inscrivendo il costrutto di intelligenza emotiva all'interno delle abilità cognitive. Dopo numerose revisioni, Mayer

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

e Salovey (1997) elaborano il noto modello a quattro rami dell'EI. In tale modello, ogni ramo rappresenta un gruppo di abilità che procede in modo evolutivo da compiti di base a compiti più complessi. In particolare, nel modello in esame i quattro rami rappresentano le seguenti abilità: percepire, valutare ed esprimere le emozioni; generare e/o utilizzare le emozioni per facilitare il pensiero; comprendere le emozioni, le loro relazioni causali, le loro trasformazioni e le combinazioni di stati emotivi; regolare e gestire le emozioni per promuovere la crescita emotiva ed intellettiva. Il modello a quattro rami dell'EI proposto da Mayer e Salovey (1997) ha subito degli aggiornamenti nel tempo (Mayer, Caruso e Salovey, 1999). Successivamente Mayer, Caruso e Salovey (2016), revisionando tale modello, hanno aggiunto nuove abilità all'interno dei quattro rami. Si veda Figura 1.

I Quattro Rami	**Tipologie di Ragionamento**
1. Percezione delle emozioni	• Identificare espressioni emotive fuorvianti o disoneste • Discriminare espressioni emotive accurate rispetto a quelle inesatte • Capire come le emozioni vengono mostrate a seconda del contesto e della cultura • Esprimere emozioni con accuratezza quando si desidera • Percepire contenuti emotivi nell'ambiente, nelle arti visive e nella musica • Percepire emozioni negli altri attraverso i loro segnali vocali, espressioni facciali, linguaggio e comportamento • Identificare emozioni nei propri stati fisici, sentimenti e pensieri

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

2. Facilitare il pensiero utilizzando l'emozione	• Selezionare problemi basandosi sul modo in cui lo stato emotivo in corso può facilitare la cognizione • Sfruttare i cambiamenti d'umore per generare diverse prospettive cognitive • Prioritizzare il pensiero dirigendo l'attenzione secondo il sentimento attuale • Generare emozioni come mezzo per relazionarsi alle esperienze di un'altra persona • Generare emozioni come supporto al giudizio e alla memoria
3. Comprensione delle emozioni	• Riconoscere le differenze culturali nella valutazione delle emozioni • Capire come una persona potrebbe sentirsi in futuro o in determinate condizioni (previsione affettiva) • Riconoscere probabili transizioni tra emozioni, come da rabbia a soddisfazione • Comprendere emozioni complesse e miste • Differenziare tra umori ed emozioni • Valutare le situazioni che probabilmente eliciteranno emozioni • Determinare antecedenti, significati e conseguenze delle emozioni • Etichettare le emozioni e riconoscere le relazioni tra di esse
4. Gestione delle emozioni	• Gestire efficacemente le emozioni degli altri per raggiungere un risultato desiderato • Gestire efficacemente le proprie emozioni per raggiungere un risultato desiderato • Valutare le strategie per mantenere, ridurre o intensificare una risposta emotiva • Monitorare le reazioni emotive per determinarne la ragionevolezza • Coinvolgere le emozioni se sono utili; disimpegnarsi se non lo sono • Restare aperti a sentimenti piacevoli e spiacevoli, se necessario, e alle informazioni che trasmettono

Figura 1. - *Il modello a quattro rami dell'Intelligenza Emotiva, con ulteriori aree di aggiornamento (Mayer, Caruso e Salovey, 2016)*

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Il costrutto di intelligenza emotiva (Mayer e Salovey, 1997) si distingue da quello di competenza emotiva (Saarni, 1990; 1999). Nel dettaglio Carolyn Saarni (1990; 1999) definisce una persona emozionalmente competente come colei in grado di utilizzare consapevolmente le proprie abilità emotive in contesti specifici, al fine di raggiungere i risultati socio-relazionali desiderati. L'autrice sottolinea, inoltre, che lo sviluppo della competenza emotiva risulta essere il frutto dell'interazione dinamica tra fattori personali (abilità di base, caratteristiche e comportamenti innati), educativi (comportamenti e abitudini apprese) e culturali (valori socialmente condivisi e accettabili) che permettono lo sviluppo di capacità morali e di giudizio. Più nello specifico Saarni (1990; 1999) articola la competenza emotiva in otto abilità legate alle emozioni: consapevolezza dei propri stati emotivi; abilità di percepire le emozioni altrui; espressione appropriata delle emozioni; empatia; capacità di distinguere tra stati emotivi manifesti ed effettivi; capacità di coping adattivo mediante autoregolazione; consapevolezza dell'importanza della comunicazione nelle relazioni interpersonali ed autoefficacia emotiva.

Zeidner e collaboratori (2003) hanno proposto un modello di comparazione tra la definizione di intelligenza emotiva di Mayer, Caruso e Salovey (1999) e la definizione di competenza emotiva proposta da Saarni (2000). Da questo confronto emerge che alcune componenti della competenza emotiva corrispondono alle abilità dell'intelligenza emotiva secondo il modello di Mayer e Salovey (1997). Infatti, le abilità che riguardano la competenza emotiva nel sé e negli altri e la distinzione tra espressione emotiva reale o fittizia corrisponderebbero in modo relativamente preciso alle abilità del primo ramo dell'intelligenza emotiva. Invece, l'abilità che riguarda l'uso del vocabolario emotivo corrisponderebbe alle abilità che compongono il terzo ramo dell'intelligenza emotiva. Infine, le modalità di autoregolazione sembrerebbero coincidere con il quarto ramo dell'intelligenza emotiva. Non trovano invece corrispondenza le abilità di empatia e le applicazioni della competenza emotiva al contesto delle relazioni. Da questo confronto emergono inoltre delle differenze tra i costrutti in esame. Nello specifico, la definizione di intelligenza emotiva di Mayer, Caruso e Salovey (1999) esclude esplicitamente la cultura, le influenze del contesto e l'autorappresentazione, mentre viene incluso il carattere morale. Inoltre, tale definizione non esamina il ruolo dello sviluppo e viene descritta essenzialmente come un costrutto che risiede internamente alla persona in quanto distinta abilità mentale. La competenza emotiva considera, invece, il contributo della storia relazionale dell'individuo (come la qualità dell'attaccamento), la complessità dello sviluppo cognitivo, il sistema di credenze e di valori in cui la persona vive ed il contesto dinamico immediato in cui le emozioni vengono evocate (Saarni, 2000). In questa sede si farà riferimento al più ampio concetto di intelligenza

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

emotiva in quanto potenziale che agisce in sinergia con le altre funzioni cognitive indipendentemente dalle influenze culturali (De Caro e D'Amico, 2008).

1.2 La relazione insegnante-studente

I contesti scolastici offrono ai bambini numerose opportunità di relazione con figure adulte diverse dalle figure genitoriali. In particolare, durante i primi anni di scuola l'insegnante può assumere il ruolo di importante figura di riferimento, sviluppando una relazione che non è solo formale, ma anche di cura. Le relazioni sperimentate a scuola forniscono agli studenti molte esperienze emotive e opportunità per apprendere abilità sociali e di autoregolazione e permettono loro di esercitare le funzioni di base dello sviluppo, come l'attaccamento, l'esplorazione, il gioco e il controllo (Pianta, 2001).

Le relazioni di attaccamento insegnante-studente hanno molte caratteristiche simili alle relazioni di attaccamento genitore-figlio e possono fornire molte delle stesse funzioni. Infatti, come le relazioni genitore-figlio, la relazione insegnante-allievo può variare in natura e qualità. In una serie di studi descrittivi, Pianta e Steinberg (Pianta e Steinberg, 1992; Pianta, 1994) hanno dimostrato che le relazioni insegnante-studente, come riportato dalla prospettiva degli insegnanti, possono essere caratterizzate dalle dimensioni di conflitto, vicinanza e dipendenza. Nello specifico il conflitto misura il grado con cui l'insegnante percepisce la sua relazione con lo studente come negativa; la vicinanza esplora in che misura la relazione insegnante-studente sia connotata da affetto, calore e comunicazione aperta; la dipendenza, invece, fa riferimento ad una eccessiva richiesta di aiuto da parte dell'alunno. Tuttavia, la qualità di queste interazioni non è definita da ciò che viene fatto singolarmente dall'adulto o dal bambino, ma piuttosto da come viene fatto in relazione all'altro (Pianta, 2001). Pertanto, costrutti come reciprocità, sensibilità, coordinazione e sincronia sono considerati importanti.

Nella letteratura internazionale sono presenti pochi studi in cui viene preso in considerazione sia il punto di vista degli insegnanti sia quello dei loro studenti in merito alla loro relazione (Elias, 1997; Di Norcia et al., 2022). Tuttavia, quando sono state confrontate le percezioni di insegnanti e studenti in merito alla loro relazione è emersa solo una debole correlazione tra le loro rispettive valutazioni rispetto alla dimensione riferita alla vicinanza (Zee e Koomen, 2017), mentre è stata rilevata un'associazione moderata tra le segnalazioni di conflitto da parte di studenti e insegnanti. Queste differenze potrebbero essere causate dalle caratteristiche, dai comportamenti e dalle convinzioni che studenti e insegnanti portano all'interno della diade. Sapere come gli alunni comprendono e valutano la relazione con

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

l'insegnante è molto importante, perché le loro percezioni possono influenzare le interazioni e lo sviluppo della relazione stessa.

Per conoscere le relazioni sociali in cui il bambino è coinvolto non è sufficiente osservare il suo comportamento, ma occorre comprendere anche cosa pensa e sente in proposito e il disegno potrebbe rivelarsi un utile alleato per questo scopo (Bombi e Pinto, 2000). Questo strumento, infatti, ha il vantaggio di essere una tecnica economica e generalmente molto gradita ai bambini. Alcuni studi, in particolare, hanno proposto il disegno come strumento di indagine della relazione insegnante-studente richiedendo ai bambini la rappresentazione di sé stessi con la propria insegnante in una situazione positiva e in una situazione negativa (Bombi et al., 2020b; Longobardi et al., 2017). Tali studi utilizzano una metodologia contrastiva con l'aspettativa che nella rappresentazione della situazione positiva i bambini rappresentino sé stessi e la propria insegnante con maggiori indicatori di benessere relazionale rispetto alla rappresentazione della situazione negativa. In questo studio per valutare il punto di vista dell'alunno è stato utilizzato lo strumento pittorico in un'ottica che sottolinea come i bambini possano essere considerati informatori competenti delle proprie esperienze (Clark e Moss, 2014).

2. Il presente studio

Nel presente studio, parte di uno studio più ampio sulla relazione tra intelligenza emotiva degli insegnati e adattamento scolastico degli alunni, ci poniamo i seguenti obiettivi specifici: confrontare il punto di vista di alunni e insegnanti sulla loro relazione; verificare il ruolo dell'EI degli insegnanti nella relazione con i propri studenti, sia nell'auto-percezione degli insegnanti sia dal punto di vista degli studenti. In particolare, saranno verificate le seguenti ipotesi. La prima ipotesi riguarda l'attesa di differenze tra le percezioni degli insegnanti e quelle dei propri studenti in merito alla loro relazione, in linea con la letteratura di riferimento (Zee e Koomen, 2017). In particolare, nel presente studio il punto di vista degli studenti è indagato attraverso l'uso del disegno, quello degli insegnanti attraverso questionari standardizzati. Ci si aspetta che l'uso di strumenti di rilevazione diversificati (disegni, questionari) possa rendere più ampie le differenze di percezione tra insegnanti e alunni. La seconda ipotesi ha l'obiettivo di verificare se maggiori livelli di EI degli insegnanti siano associati a migliori relazioni con gli alunni. In particolare, si ipotizza che nelle classi in cui i docenti presentano livelli più elevati di EI le relazioni insegnanti-studenti siano caratterizzate da maggiori livelli di vicinanza e minori livelli di conflitto e dipendenza. Infatti, la letteratura riferisce che gli insegnanti dotati di livelli più elevati di EI

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

affrontano le situazioni conflittuali in modo più costruttivo rispetto ai colleghi con livelli di EI inferiori (Jeloudar et al., 2011; Perry e Ball, 2007; Ramana, 2013). Inoltre, insegnanti con livelli più alti di EI costruiscono relazioni positive con i loro studenti e affrontano efficacemente le difficoltà comportamentali di questi ultimi (Brackett et al., 2011; Nizielski et al., 2012). In particolare, lo studio di Poulou (2017) mette in evidenza che i livelli di EI mostrano una correlazione positiva con la dimensione della vicinanza nella relazione insegnante-studente. Inoltre, l'autonomia si sviluppa all'interno di relazioni adulto-bambino positive, in cui l'adulto è in grado di sostenere e tollerare l'autonomia del bambino (Pianta, 2001).

3. Metodo

3.1 Partecipanti

Hanno partecipato allo studio in totale 79 alunni di scuola primaria, 41 di classe terza (distribuiti in tre sezioni) e 38 di classe quarta (distribuiti in tre sezioni). I partecipanti sono 48 bambini e 31 bambine, di età compresa tra gli otto e gli undici anni (età media 8,67; DS=.59), le loro insegnanti prevalenti e i loro genitori. Il 4,7% degli alunni che hanno partecipato alla ricerca proviene da un contesto migratorio. I genitori rispondenti sono per l'80% madri. Per quel che riguarda il titolo di studio, il 10% dei genitori partecipanti ha diploma di scuola secondaria di primo grado, il 38% di scuola secondaria di secondo grado, il 46% laurea di primo o secondo livello e il rimanente 6% ha un titolo di studio post-laurea. Le insegnanti sono tutte donne, una ha meno di 30 anni, 2 hanno un'età compresa tra 40 e 50 anni e 3 tra 50 e 60 anni, gli anni di insegnamento vanno dai 6 ai 28.

3.2 Procedura

La ricerca è stata condotta in una scuola primaria di una città del Lazio in cui risiede l'8,3 per cento della popolazione regionale. Il 23,1% degli abitanti della cittadina ha un'età superiore ai 65 anni e gli studenti iscritti a scuola sono complessivamente 69.049, di cui 21.044 sono iscritti alla scuola primaria. La scuola fa parte di un istituto comprensivo che include complessivamente quattro scuole dell'infanzia, tre scuole primarie e una scuola secondaria di primo grado per un totale di 716 alunni e studenti, divisi in 47 classi. L'istituto è situato nel centro storico della città ed è frequentato da alunni provenienti dai quartieri del centro storico, dalla campagna limitrofa e da alcuni comuni circostanti. È in aumento la frequenza di alunni provenienti

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

da contesti migratori, in particolare di origini rumene, albanesi, ucraine e tunisine. La scuola primaria in cui si è svolta la ricerca ospita 280 alunni dai 5 agli 11 anni divisi in 15 classi. Il plesso dispone di undici aule, un'aula multimediale attrezzata, un laboratorio di informatica, un'aula per il sostegno ed una palestra. Le attività didattiche si svolgono dal lunedì al sabato, iniziano alle 8:00 e si concludono alle 13:10.

Dopo aver ricevuto il permesso dalla scuola di svolgere lo studio, sono stati presi accordi per consegnare i consensi informati e i questionari ai genitori degli studenti. Una volta ottenuti i consensi necessari, la somministrazione degli strumenti è avvenuta in due mattinate, dalle 9:00 alle 13:00. La somministrazione non è avvenuta in presenza delle insegnanti di classe. In aula, abbiamo inizialmente presentato lo studio agli alunni e risposto alle loro domande. La somministrazione è iniziata distribuendo i questionari[1] agli alunni i cui genitori avevano consegnato il consenso informato. Gli altri alunni hanno svolto altre attività. Dopo aver illustrato la consegna e averla scritta alla lavagna, sono stati consegnati, ad un alunno per volta, i fogli su cui disegnare. La consegna è stata di disegnare, utilizzando solo matita e gomma, sé stessi e la propria insegnante prima in un momento in cui le cose andavano bene (armonia) e poi in un momento cui le cose andavano meno bene (disarmonia). Al termine della somministrazione è stato previsto del tempo per rispondere a eventuali dubbi o richieste degli alunni. I questionari per le insegnanti sono stati consegnati alle partecipanti prima di entrare in classe.

3.3 Strumenti

Alle insegnanti sono stati somministrati i seguenti strumenti:

- La Student-Teacher Relationship Scale, uno strumento, costruito da Pianta (1994) standardizzato e validato per valutare la percezione che hanno gli insegnanti della relazione con i propri studenti. Nella sua versione originale validata su popolazioni nordamericane, la scala, composta da 26 items, consente di misurare tre dimensioni della qualità della relazione: il conflitto, la vicinanza affettiva e la dipendenza. Il conflitto misura il grado con cui l'insegnante percepisce la sua relazione con lo studente come negativa; la vicinanza esplora in che misura la relazione insegnante-studente sia connotata da affetto, calore e comunicazione aperta; la dipendenza invece fa riferimento a una eccessiva richiesta di aiuto da parte dell'alunno. Lo strumento viene compilato dall'insegnante

[1] La ricerca più ampia ha previsto la somministrazione agli alunni di questionari sulla prosocialità i cui risultati non sono oggetto del presente studio.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

in riferimento a ogni singolo alunno della sua classe, al fine di individuare quegli studenti a rischio dal punto di vista dei comportamenti, dei rapporti sociali, delle future traiettorie di sviluppo e dell'adattamento scolastico. La modalità di risposta è su una scala a sei passi (da 0= per niente d'accordo a 5=del tutto d'accordo).

- La Self-Rated Emotional Intelligence Scale (SREIS; Brackett, Rivers, Shiffman, Lerner, e Salovey, 2006) è stata utilizzata per misurare la percezione degli insegnanti del proprio livello di EI. In particolare, la scala è costituita da 33 items con modalità di risposta su una scala a 5 passi (da 0=per niente d'accordo a 4=del tutto d'accordo). La Self-Rated Emotional Intelligence Scale si basa sul modello di intelligenza emotiva di Mayer e Salovey (1990) e comprende:
    1. una sottoscala di percezione delle emozioni: 4 item (ad esempio "riconosco le emozioni che le persone provano osservando le loro espressioni facciali");
    2. una sottoscala di utilizzo delle emozioni: 3 item (ad esempio "quando prendo decisioni, ascolto i miei sentimenti per decidere se la decisione è giusta");
    3. una sottoscala di comprensione delle emozioni: 4 item (ad esempio "ho un vocabolario ricco per descrivere le mie emozioni");
    4. una sottoscala di gestione delle emozioni: 8 item di cui 4 relativi alla gestione delle proprie emozioni (ad esempio, "ho problemi a gestire i miei sentimenti di rabbia") e 4 relativi alla gestione sociale delle emozioni (ad esempio, "ho strategie per migliorare l'umore degli altri").

Agli alunni è stato chiesto di fare due disegni, utilizzando solo la gomma e la matita, in particolare, uno di una situazione relazionale positiva (armonia) con la loro insegnante e uno in una negativa (disarmonia).

Per la codifica dei disegni sono state utilizzate le scale di coesione e di distanziamento del metodo PAIR (Bombi et al., 2007). In particolare, ciascuna scala include sei sottoscale codificate dicotomicamente (0=assenza; 1=presenza di uno o più indici pittorici). Ogni subscala di coesione ha una corrispondente subscala di distanziamento. Di seguito sono riportate le scale prese in esame con le relative subscale:
1. Coesione: la coesione tra le figure si misura con sei distinte subscale: (C1) sguardo; (C2) avvicinamento; (C3) attività coordinata; (C4) vicinanza; (C5) area comune; (C6) unione. 2. Distanziamento: il distanziamento tra le figure si misura con sei distinte subscale: (D1) sguardo distolto; (D2) allontanamento; (D3) attività indipendente; (D4) lontananza; (D5) area individuale; (D6) separazione.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Non si tratta di due dimensioni polarizzate, ma di un continuum, tanto che le dimensioni possono coesistere nello stesso disegno. Negli studi più recenti sulla relazione insegnante-studente (Bombi et al., 2020a) la dimensione della coesione sembra essere una misura della dipendenza o dell'insicurezza, mentre il distanziamento è una misura dell'autonomia individuale.

3.4 Analisi

I dati sono stati analizzati utilizzando il pacchetto SPSS Statistics 27. Dopo aver illustrato le statistiche descrittive, i punteggi EI sono stati utilizzati per suddividere le sei insegnanti in tre gruppi, con livelli di EI bassi (EI<2.5), medi (2.5<EI<3.00) e alti (EI>3.00). I tre gruppi sono stati utilizzati come variabili indipendenti in una serie di ANOVA aventi dimensioni di STRS e PAIR per variabili dipendenti.

4. Risultati

4.1 Punti di vista a confronto su relazione insegnante alunno

	N	Minimo	Masimo	Media	Deviazione standard
EI Comprensione emozioni altrui	6	2,11	2,78	2,50	,24
EI Comprensione emozioni proprie	6	2,00	3,50	2,50	,61
EI Regolazione emotiva	6	2,67	3,44	3,11	,28
Intelligenza Emotiva Totale	6	2,33	3,09	2,70	,32
PAIR Coesione (Armonia)	79	0	4	1,16	,93

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

PAIR Distanziamento (Armonia)	79	0	3	1,05	1,10
PAIR Coesione (Disarmonia)	78	0	4	1,29	,81
PAIR Distanziamento (Disarmonia)	78	0	4	1,40	1,20
STRS Conflitto	79	1,33	3,00	1,91	,31
STRS Vicinanza	79	1,22	3,22	2,08	,32
STRS Dipendenza	79	2,00	4,80	3,39	,58

Tab. 1 - Statistiche Descrittive

La tabella 1 mostra le statistiche descrittive, mentre la tabella 2, riportata di seguito, mostra le correlazioni ottenute confrontando i punteggi ottenuti dall'STRS nelle dimensioni di conflitto, vicinanza e dipendenza con i punteggi nelle scale di coesione e distanziamento del disegno, codificato con il metodo PAIR. In particolare, le analisi mostrano una correlazione moderata (r=.267) tra la scala di coesione del disegno e la dimensione della dipendenza dell'STRS. Non emergono altre correlazioni significative.

	STRS Conflitto	STRS Vicinanza	STRS Dipendenza
PAIR Coesione (armonia)	,165	,174	,267*
PAIR Distanziamento (armonia)	,030	,187	,075

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

PAIR Coesione (disarmonia)	,047	,160	,031
PAIR Distanziamento (Disarmonia)	-,119	,001	-,020

*<.05

Tab. 2 - *confronto STRS e disegno*

4.2 Relazione insegnante-studente e intelligenza emotiva secondo la percezione delle insegnanti

La tabella 3 mostra le medie dei punteggi ottenuti dall'STRS nelle tre dimensioni: conflitto, vicinanza e dipendenza nei tre gruppi: intelligenza emotiva bassa, intelligenza emotiva media, intelligenza emotiva alta e il punteggio medio totale.

	N	Minimo	Massimo	Deviazione standard
STRS Conflitto	EI Bassa	29	1,93	,23
	EI Media	20	1,95	,44
	EI Alta	30	1,86	,28
	Totale	79	1,91	,31
STRS Vicinanza	EI Bassa	29	2,06	,31
	EI Media	20	2,20	,29
	EI Alta	30	2,02	,33
	Totale	79	2,08	,32

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

STRS Dipendenza	EI Bassa	29	3,69	,50
	EI Media	20	3,16	,50
	EI Alta	30	3,25	,59
	Totale	79	3,39	,58

Tab. 3 - *STRS e Intelligenza Emotiva*

Le ANOVA eseguite sulle dimensioni della relazione insegnante-alunno, con i gruppi di EI per VI, mostrano una differenza significativa per la dimensione della dipendenza ($F_{2,78}=7,36$; $p=.001$) nel gruppo in cui le insegnanti hanno livelli di EI più bassi.

Di seguito sono illustrati, nelle figure 2, 3 e 4, i punteggi riportati nella tabella 3.

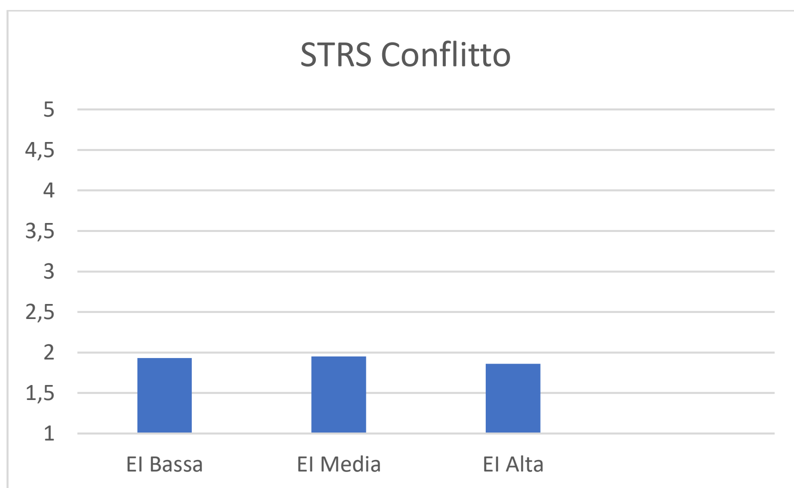

Figura 2 - *STRS Conflitto e Intelligenza Emotiva*

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

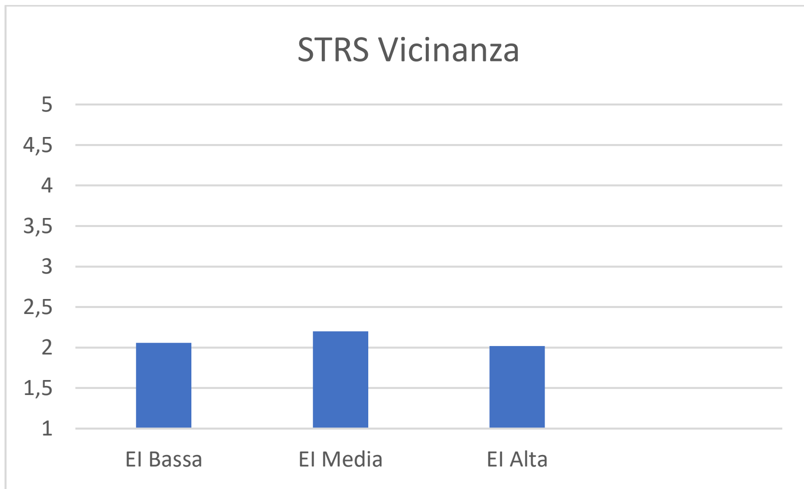

Figura 3 - *STRS Vicinanza e Intelligenza Emotiva*

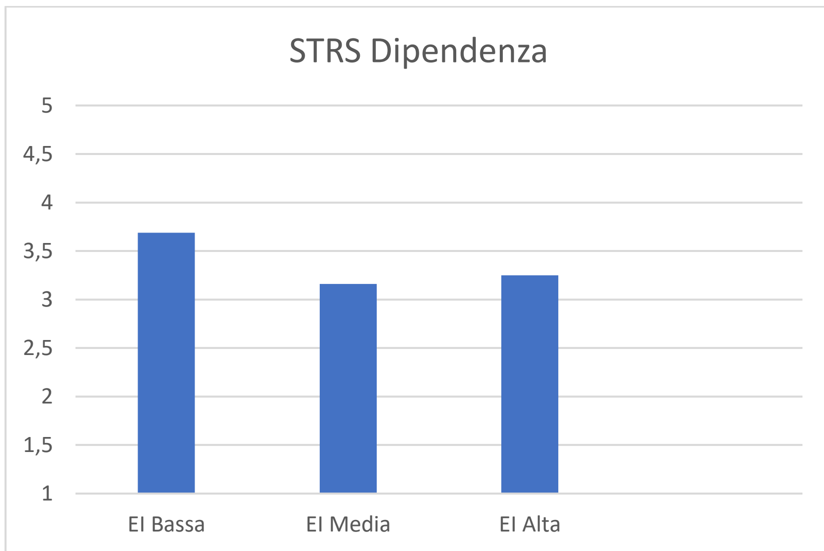

Figura 4 - *STRS Dipendenza e Intelligenza Emotiva*

Come rappresentato dalle figure 2, 3 e 4, il gruppo di insegnanti con i livelli di EI più bassi percepiscono maggiore dipendenza nella relazione con i loro studenti. In particolare, la dimensione della dipendenza fa riferimento ad una eccessiva richiesta di aiuto da parte dell'alunno e può essere

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

considerata come un indicatore del basso livello di autonomia del bambino (Pianta, 2001).

4.3 Relazione insegnante-studente e intelligenza emotiva secondo la percezione degli studenti

La tabella 4 mostra i risultati delle ANOVA eseguite sulle variabili del disegno.

		N	Medio	Deviazione standard
PAIR Coesione (disarmonia)	EI Bassa	29	1,31	,76
	EI Media	19	1,16	,96
	EI Alta	30	1,37	,77
	Totale	78	1,29	,81
PAIR Distanzia-mento (disarmonia)	EI Bassa	29	1,41	1,30
	EI Media	19	1,00	1,00
	EI Alta	30	1,63	1,19
	Totale	78	1,40	1,20
PAIR Coesione (armonia)	EI Bassa	29	1,45	1,02
	EI Media	20	1,15	,81
	EI Alta	30	,90	,85
	Totale	79	1,16	,93
PAIR distanzia-mento (armonia)	EI Bassa	29	,93	1,07
	EI Media	20	,80	1,06
	EI Alta	30	1,33	1,12
	Totale	79	1,05	1,10

Tab. 4 - *Intelligenza Emotiva e Disegno*

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Le ANOVA eseguite sulle variabili del disegno mostrano una differenza tendenziale per la dimensione della coesione nel disegno delle situazioni di armonia ($F_{2,78}=1,7$; $p=.07$); nelle classi in cui le insegnanti hanno livelli di EI più bassi, gli alunni rappresentano più coesione. In particolare, gli alunni appartenenti alle classi con le insegnanti dotate di livelli di EI inferiore mostrano, nei disegni raffiguranti momenti di relazione positiva, maggiori indicatori pittorici per la dimensione della coesione, a differenza dei bambini appartenenti alle classi con insegnanti dotate di livelli di EI medi ed alti.

5. Discussione e conclusioni

Lo studio presentato ha lo scopo di indagare il ruolo dell'intelligenza emotiva negli insegnanti di scuola primaria nella qualità della relazione insegnante-studente.

In questo studio, in coerenza con la letteratura, sono state formulate due ipotesi. In primo luogo, è stato ipotizzato che confrontando il punto di vista di insegnanti e studenti sulla loro relazione sarebbero emerse delle differenze tra le percezioni. Questa ipotesi è stata confermata dalle correlazioni ottenute confrontando i punteggi dell'STRS nelle dimensioni di conflitto, vicinanza e dipendenza con i punteggi nelle scale di coesione e distanziamento del disegno codificato con il metodo PAIR: si osserva una correlazione moderata ($r=.26$) tra la scala di coesione del disegno e la dimensione della dipendenza dell'STRS. Questa correlazione è in linea con quanto presente in letteratura; infatti, la dimensione della coesione, anche in altri studi, correla con il bisogno degli alunni di rimanere vicini alle insegnanti e con l'insicurezza (Bombi et al., 2020a).

In secondo luogo, è stato ipotizzato che maggiori livelli di EI nelle insegnanti si potessero presentare insieme a migliori relazioni con gli studenti. In particolare, ci si aspettava che nelle classi in cui le docenti presentavano maggiori livelli di EI le relazioni insegnanti-studenti sarebbero risultate caratterizzate da maggiori livelli di vicinanza e minori livelli di conflitto e dipendenza. In particolare, i punteggi di EI sono stati utilizzati per suddividere le sei insegnanti in tre gruppi, con livelli di EI bassi, medi e alti. I tre gruppi sono stati confrontati rispetto ai dati ottenuti da STRS e PAIR. I risultati ottenuti mostrano che la dimensione della dipendenza è risultata significativamente più elevata nel gruppo le cui insegnanti hanno livelli di EI più bassi. Dunque, rispetto a questa dimensione le insegnanti con livelli di EI medio/alti presentano relazioni con i loro studenti migliori, rispetto alle colleghe dotate di livelli di EI inferiori.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

I risultati ottenuti sono in linea con quanto presente in letteratura. Infatti, l'autonomia è costruita sulle basi relazionali dell'attaccamento ed è supportata nel contesto delle relazioni attualmente disponibili al bambino e dal modo in cui tali relazioni tollerano e sostengono l'autonomia dell'individuo e allo stesso tempo forniscono una base sicura e un sostegno sostanziale (Pianta, 2001). In particolare, gli insegnanti con livelli di EI elevata tendono a costruire relazioni positive con i loro studenti (Brackett et al., 2011; Nizielski et al., 2012) e nelle relazioni positive l'adulto mostra maggiori capacità nel leggere accuratamente i segnali del bambino, nel rispondere in modo contingente sulla base di questi segnali, nel trasmettere senso di accoglienza e calore emotivo, nell'offrire l'assistenza necessaria, nel modellare un comportamento e nel mettere in atto strutture e limiti (Howes et al., 1994). Queste capacità permettono all'adulto di fornire la base sicura e il sostegno necessario per lo sviluppo dell'autonomia del bambino (Pianta, 2001). Al contrario, l'incapacità nella regolazione delle emozioni, un indicatore dell'EI secondo il modello a quattro rami di Mayer e Salovey (1990), può portare l'insegnante a un eccessivo coinvolgimento e a un'eccessiva preoccupazione verso l'altro, producendo nei bambini eccessivi livelli di ansia e dipendenza (Pianta, 2001).

Le analisi eseguite sulle variabili del disegno mostrano che nelle classi i cui insegnanti hanno livelli di EI più bassi, i bambini e le bambine rappresentano più coesione. La coesione è un indicatore della dipendenza e dell'insicurezza (Bombi et al., 2020a), quindi in questo caso le insegnanti con livelli di EI medio/alti presentano relazioni con i loro studenti meno dipendenti e meno insicure rispetto alle colleghe con livelli di EI più bassi. La dipendenza dei bambini nei confronti delle loro insegnanti è correlata ad una ricerca di prossimità fisica da parte del bambino (Pianta, 2001). Un'eccessiva coesione, rappresentata dai bambini nelle situazioni di relazione positiva, verso le insegnanti con livelli di EI inferiori, potrebbe fare riferimento ad un'eccessiva dipendenza nei confronti dell'insegnante.

Non sono state riscontrate differenze significative, tra i tre gruppi di insegnanti, in merito alle dimensioni di vicinanza e conflitto, come invece era stato ipotizzato. Questo aspetto costituisce un limite dello studio che potrà essere approfondito in successive ricerche. Infatti, questa ricerca è stata condotta su un ridotto numero di partecipanti ed è attualmente in corso un ampliamento del campione. Inoltre, sarebbe utile indagare ulteriormente la percezione della relazione da parte degli alunni attraverso questionari, i cui risultati potrebbero integrare i dati rilevati attraverso il disegno.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

5.1 Implicazioni

Per le ricerche future, il disegno si è rivelato un utile strumento per indagare le percezioni degli alunni in merito alla relazione con le proprie insegnanti dal momento che è un compito ecologico, facile da eseguire in classe (Bombi et al., 2020a).

Una delle implicazioni del presente studio, potrebbe essere riflettere sul tema della formazione pre-servizio e in servizio degli insegnanti.

I programmi di intervento orientati alla promozione e allo sviluppo dell'intelligenza emotiva e delle competenze socio-emotive sono numerosi, ma prevalentemente destinati ad alunni e studenti di qualsiasi fascia di età, dalla prima infanzia all'età adulta (Puertas Molero et al., 2020) oppure a figure professionali manageriali o con posizioni di leadership (Caruso e Salovey, 2004; Higgs e Dulewicz, 2024).

I contesti educativi sono contesti organizzativi a elevata complessità, poiché richiedono lo sviluppo e il continuo aggiornamento di un gran numero di conoscenze, competenze e abilità nel corso di tutta la carriera professionale, un adattamento ai cambiamenti continui della società e della cultura, una gestione equilibrata delle relazioni con una molteplicità di attori (alunni/studenti, genitori/tutori legali, colleghi, dirigenti scolastici, personale tecnico amministrativo). Per tale motivo, la letteratura sul rischio burnout della professionalità docente è corposa (García-Carmona et al., 2019) e la ricerca sulla promozione dell'intelligenza emotiva degli insegnanti attraverso la formazione si focalizza prevalentemente proprio sulla funzione che potrebbe svolgere nella diminuzione dei livelli di stress di lavoro correlato e di burnout (Oliveira et al., 2021).

A partire dal presente studio, potrebbe essere interessante condurre delle ricerche longitudinali, durante tutta la durata della scuola primaria, per valutare come la dimensione della dipendenza dell'STRS muta durante lo sviluppo, nei gruppi con insegnanti con livelli di EI bassi e medio/alti. Conoscere meglio il ruolo dell'EI nella qualità della relazione insegnante-alunno potrebbe suggerire e supportare la progettazione di interventi di formazione continua degli insegnanti, volti al potenziamento dell'intelligenza emotiva con l'obiettivo di supportare la costruzione di relazioni positive tra insegnanti e alunni e, dunque, di migliorare il clima di classe e i livelli di benessere di insegnanti e alunni e di potenziare la qualità dei processi di apprendimento-insegnamento.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Riferimenti bibliografici

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology*, *35*(1), 61-79. DOI: 10.1016/S0022-4405(96)00029-5.

Birch, S. H., & Ladd, G. W. (1998). Children's interpersonal behaviors and the teacher-child relationship. *Developmental Psychology, 34*(5), 934-946. DOI: 10.1037/0012-1649.34.5.934.

Bombi, A. S., & Pinto, G. (2000). *Le relazioni interpersonali del bambino. Studiare la socialità infantile con il disegno*. Roma: Carocci.

Bombi, A. S., Pinto, G., & Cannoni, E. (2007). *Pictorial assessment of interpersonal relationships (PAIR)*. Firenze: University Press.

Bombi, A. S., Cannoni, E., & Galli, F. (2020a). The relationship between teachers and students: Children' pictorial representation and teachers' evaluation. *Current Psychology, 5*(7), 43-57. DOI: 10.1007/s12144-020-00702-7.

Bombi, A. S., Cannoni, E., Gallì, F., & Di Norcia, A. (2020b). Students' and teachers' emotions. A study with children's drawings. *Psychology Hub*, *37*(1), 13-18. DOI: 10.13133/2724-2943/16897.

Brackett, M. A., Rivers, S. E., Shiffman, S., Lerner, N., & Salovey, P. (2006). Relating emotional abilities to social functioning: A comparison of self-report and performance measures of Emotional Intelligence. *Journal of Personality and Social Psychology*, *91*(4), 780-795. DOI: 10.1037/0022-3514.91.4.780.

Brackett, M. A., Reyes, M. R., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2011). Classroom emotional climate, teacher affiliation and student conduct. *Journal of Classroom Interaction*, *46*(1), 27-36. DOI: 10.1037/a0027268.

Caruso, D. R., & Salovey, P. (2004). *The Emotionally Intelligent Manager: How to Develop and Use the Four Key Emotional Skills of Leadership*. San Francisco, Calif.: Jossey-Bass.

Cefai, C., & Cavioni, V. (2014). *Social and emotional education in primary school. Integrating theory and research into practice*. New York: Springer.

Clark, A., & Moss, P. (2014). *Ascoltare i bambini. L'approccio a mosaico*. Junior-Spaggiari.

De Caro, T., & D'Amico, A. (2008). L'intelligenza emotiva: rassegna dei principali modelli teorici, degli strumenti di valutazione e dei primi risultati di ricerca. *Giornale italiano di psicologia*, *35*(4), 857-884. DOI: 10.1421/28421.

Di Norcia, A., Bombi, A.S., Pinto, G., & Cannoni, E. (2022). The student's drawing of teacher's pictorial Value as a predictor of the student-teacher relationship and school adjustment. *Frontiers in Psychology, 13*, 1006568. DOI: 10.3389/fpsyg.2022.1006568.

Downer, J. T., Rimm-Kaufman, S. E., & Pianta, R. C. (2007). How do classroom conditions and children's risk for school problems contribute to children's behavioral engagement in learning?. *School Psychology Review, 36*(3), 413-432.

Elias, M. (1997). The Missing Piece: Making the Case for Greater Attention to Social and Emotional Learning in the Schools. *Education Week*, *17*(5), 36-38.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

García-Carmona, M., Marín, M.D. & Aguayo, R. (2019). Burnout syndrome in secondary school teachers: a systematic review and meta-analysis. *Soc Psychol Educ, 22*, 189-208. DOI: 10.1007/s11218-018-9471-9.

Gómez-Leal, R., Holzer, A. A., Bradley, C., Fernández-Berrocal, P., & Patti, J. (2022). The relationship between emotional intelligence and leadership in school leaders: a systematic review. *Cambridge Journal of Education*, *52*(1), 1-21. DOI: 10.1080/0305764X.2021.1927987.

Gunter, L., Caldarella, P., Korth, B. B., & Young, K. R. (2012). Promoting social and emotional learning in preschool students: A study of Strong Start Pre-K. *Early Childhood Education Journal, 40,* 151-159. DOI: 10.1007/s10643-012-0507-z.

Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development, 72*(2), 625-638. DOI: 10.1111/1467-8624.00301.

Howes, C., Matheson, C. C., & Hamilton, C. E. (1994). Maternal, teacher, and child care history correlates of children's relationships with peers. *Child development*, *65*(1), 264-273. DOI: 10.1111/j.1467-8624.1994.tb00749.x.

Higgs, M., & Dulewicz, V. (2024). *Leading with emotional intelligence: Effective change implementation in today's complex context*. Cham: Palgrave Macmillan.

Jeloudar, S. Y., Yunus, A. S. M., Roslan, S., & Nor, S.M. (2011). Teachers Emotional Intelligence and its relationship with classroom discipline strategies based on teachers and student's perceptions. *Journal of Psychology, 2*(2), 95-102. DOI: 10.1080/09764224.2011.11885468.

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of educational research*, *79*(1), 491-525. DOI: 10.3102/0034654308325693.

Longobardi, C., Pasta, T., Gastaldi, F. G. M., & Prino, L. E. (2017/05//). Measuring the student-teacher relationship using children's drawings in an Italian elementary school. *Journal of Psychological and Educational Research, 25*(1), 115-129.

Madigan, D. J., Kim, L. E., Glandorf, H. L., & Kavanagh, O. (2023). Teacher burnout and physical health: A systematic review. *International Journal of Educational Research*, *119*, 102173. DOI: 10.1016/j.ijer.2023.102173.

Mancini, G., Mameli, C. & Biolcati, R. (2022). Burnout in Italian Primary Teachers: The Predictive Effects of Trait Emotional Intelligence, Trait Anxiety, and Job Instability. *Eur J Psychol.*, *18*(2), 168-180. DOI: 10.5964/ejop.2685.

Mayer, J. D, & Salovey P. (1990). Emotional intelligence. *Imagination, cognition and personality*, *9*(3), 185-211. DOI: 10.2190/DUGG-P24E-52WK-6CDG.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey e D. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications*. New York: Basic Books.

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, *27*(4), 267-298. DOI: 10.1016/S0160-2896(99)00016-1.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Mayer, J. D., Caruso, D. R. & Salovey, P. (2016). The ability model of emotional intelligence: Principles and updates. *Emotion review*, *8*(4), 290-300. DOI: 10.1177/1754073916639667.

Meehan, B. T. (2004). Teacher-student relationships as compensatory resources for aggressive children. *The Sciences and Engineering*, *64*, 40-51.

Molinari, L., & Melotti, G. (2010). La relazione fra insegnanti e alunni nella scuola primaria: un contributo alla validazione italiana della Student-Teacher Relationship Scale (STRS). *Rassegna di Psicologia*, *27*(2), 9-34.

Nizielski, S., Hallum, S., Lopes, P. N. & Schutz, A. (2012). Attention to student needs mediates the relationship between emotional intelligence and student misconduct in the classroom. *Journal of Psychoeducational Assessment, 30*(4), 320-329. DOI: 10.1177/0734282912449439.

Oliveira, S., Roberto, M.S., Veiga-Simão, A.M, Marques-Pinto, A. (2021). A Meta-analysis of the Impact of Social and Emotional Learning Interventions on Teachers' Burnout Symptoms. *Educ Psychol Rev,* 33, 1779-1808. DOI: 10.1007/s10648-021-09612-x.

Perry, C., & Ball, I. (2007). Dealing constructively with negative evaluated emotional situations: the key to understanding the different reactions of teachers with high and low levels of emotional intelligence. *Social Psychology of Education, 10,* 443-454. DOI: 10.1007/s11218-007-9025-z.

Pianta, R. C., & Steinberg, M. (1992). Relationships between children and kindergarten teachers from the teachers' perspective. In R. C. Pianta (Ed.), *Beyond the parent: The role of the adults in children's lives* (pp. 61-80). San Francisco: Jossey-Bass.

Pianta, R. C. (1994). Patterns of relationships between children and kindergarten teachers. *Journal* of *School Psychology, 32,* 15-32. DOI: 10.1016/0022-4405(94)90026-4.

Pianta, R. C. (2001). *La relazione bambino-insegnante. Aspetti evolutivi e clinici [Enhancing Relationships Between Children and Teachers].* Milano: Raffaello Cortina.

Poulou, M. S. (2017). Students' emotional and behavioral difficulties: The role of teachers' social and emotional learning and teacher-student relationships. *International Journal of Emotional Education, 9*(2), 72-89.

Puertas Molero, P., Zurita Ortega, F., Ubago Jiménez, J. L. & González Valero, G. (2019). Influence of Emotional Intelligence and Burnout Syndrome on Teachers Well-Being: A Systematic Review. *Soc. Sci.*, *8*, 185. DOI: 10.3390/socsci8060185.

Puertas Molero, P., Zurita-Ortega, F., Chacon-Cuberos, R., Castro-Sanchez, M., Ramirez-Granizo, I., & Gabriel, G. V. (2020). Emotional intelligence in the educational field: A meta-analysis. *Anales De Psicología, 36*(1), 84-91. DOI: 10.6018/analesps.345901.

Ramana, T. V. (2013). Emotional intelligence and teacher effectiveness – An analysis. *Voice of Research, 2*(2), 18-22.

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The Influence of Affective Teacher-Student Relationships on Students' School Engagement

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

and Achievement: A Meta-Analytic Approach. *Review of Educational Research*, *81*(4), 493-529. DOI: 10.3102/0034654311421793.

Saarni, C. (1990). Emotional competence: How emotions and relationship become integrated. In R.A. Thompson (ed.), *Socioemotional development. Nebraska symposium on motivation* (pp. 115-182). Lincoln, NE: University of NebraskaPress.

Saarni, C. (1999). *The development of emotional competence*. New York: Guilford Press.

Saarni, C. (2000). *The Social Context of Emotional Development*. In M. Lewis, J. Haviland (Eds.), *The Handbook of Emotion* (pp. 306-322). New York: Guilford Press.

Valente, S., Veiga-Branco, A., Rebelo, H., Lourenço, A. A. & Cristóvão, A. M. (2020). The Relationship between Emotional Intelligence Ability and Teacher Efficacy. *Universal Journal of Educational Research*, *8*(3): 916-923. DOI: 10.13189/ujer.2020.080324.

Zee, M., & Koomen, H. M. (2017). Similarities and dissimilarities between teachers' and students' relationship views in upper elementary school: The role of personal teacher and student attributes. *Journal of School Psychology*, *64*, 43-60. DOI: 10.1016/j.jsp.2017.04.007.

Zeidner, M. (2003). Development of Emotional Intelligence: Towards a Multilevel Investment Model. *Human Development*, *46*(2), 69-96. DOI: 10.1159/000068580.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

The relationship between theory of mind and emotion regulation in the preschool years. The role of language and gender

Daniela Bulgarelli°*, Maria Nives Sala°**, Beatrice Ragaglia°***, Paola Molina[§]

°Department of Psychology,
Università degli Studi di Torino,
Torino, Italia;
*e-mail: daniela.bulgarelli@unito.it.
** e-mail: nives.sala@gmail.com.
*** e-mail: beatrice.ragaglia@unito.it.
[§]Interuniversity Department of Regional and Urban Studies and Planning,
Università degli Studi di Torino,
Torino, Italia;
e-mail: paola.molina@unito.it.

Ricevuto: 08.05.2024. - **Accettato**: 15.07.2024

Pubblicato online: 23.12.2024

Abstract

A few studies investigated the relationship between Theory of Mind (ToM) and Emotion Regulation (ER) in preschoolers, mainly using False Belief tasks for assessing ToM, and disappointment procedures for measuring ER. Mixed findings were reported, as some researchers did not find any association and others did. The current study aims at testing the positive correlation between ToM and ER, using different measures, i.e., the comprehensive test ToM Storybooks, which allows for qualitative, quantitative, and total scores, and the narrative story stem procedure SIRE, which returns four scores: Behavioural Strategy, Cognitive Reappraisal, Social Support, and Attentional Deployment. Moreover, the study aims at testing the role of receptive vocabulary size (VS) and gender on the relationship between ToM and ER. Sixty-two Italian 3- to 6-year-old children participated. The three ToM scores, VS score, and ER Behavioural Strategy and Cognitive Reappraisal scores were associated with age. Gender differences were only found in ER Behavioural Strategy and Social Support scores. VS was significantly associated with the three ToM scores and with the ER Behavioural Strategy and Cognitive Reappraisal.

ISSNe 1972-5620, Doi: 10.3280/rip2024oa18947

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Partial correlations controlling for age, gender, and VS showed that ToM Qualitative score was associated with ER Cognitive Reappraisal: more complex ToM abilities were associated with ER cognitive strategies.

Keywords: ToM Storybooks; Emotion Regulation Story Stems; Social Cognition; Receptive Vocabulary Size; Gender differences.

Introduction

The current study aims at analysing the relationships among Theory of Mind (ToM), Emotion Regulation (ER), language, and gender in typically developing preschoolers. In what follows, the literature about the associations among these competencies is presented. As discussed below, only five studies from 2007 and 2015 analysed together the four skills.

Theory of Mind, language, and gender

ToM is the ability to attribute mental states, such as desires, thoughts, intentions, and beliefs to self and others; to understand that others' mental states differ from one's own; and to predict one's own and others' behaviour (Wellman, 2018). ToM develops from early infancy: starting from 2 years of age, children talk about emotional experiences, and understand that desires cause behaviours and emotions; at about 3 years, they use beliefs to predict behaviours and, from 4 years onwards, they understand that mental states can differ from reality and be untrue, and children can use false beliefs to foretell the other people's behaviours (Blijd-Hoogewys et al., 2022; Wellman, 2018).

Since the early Eighties, False Belief tasks have been mainly used to assess ToM abilities in preschool age (Wellman, 2018). Several versions of the False Belief tasks have been developed, but all are characterized by asking the child to predict the behaviour of another person, based on the fact that this person has an erroneous belief in reality. The children are told a story or observe some puppets interacting with each other, and have to solve such problems: where Ann will look for a doll if she thinks it is in a basket, even though it is actually in a box? What will Maxi expect to find in a closed box of Smarties that instead contains pencils? The developmental trend of children's responses to the False Belief procedures does not depend on the type of task, type of questions, nature of the protagonist, etc. (Wellman, 2018). Researchers grew the awareness that ToM is not simply understanding False Beliefs and several instruments have been designed to better analyse the

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

construct. A systematic review reported that, from 1983 to 2019, 220 different measures have been used to assess ToM in children from 0 to 5 years of age (Beaudoin et al., 2020). These tasks and tools measure abilities that could be grouped into seven clusters of psychological aspects that contribute to the multifaceted construct of ToM: emotions, desires, intentions, percepts, knowledge, beliefs, and mentalistic understanding of non-literal communication. Each cluster included from four to eight sub-abilities. To obtain a more reliable and complex assessment, currently 18 comprehensive measures, composed of multiple ToM tasks, are available. The ToM Storybooks is one of them (Blijd-Hoogewys et al., 2008; Bulgarelli et al., 2015). In this tool, five ToM components are evaluated: psychology of emotions, desires, beliefs, perception leads to knowledge, and distinction between mental and physical phenomena. The components are presented via two or more similar tasks contextualised in different stories, thus allowing for more reliable and stable measurement than FB tasks alone. A detailed description of the test is reported in the Supplementary materials.

ToM and language are interdependent competences (Milligan et al., 2007): longitudinal studies showed that significant predictive correlations between language and ToM were greater in magnitude than those between ToM and language. The relationship between ToM and several linguistic and communication features – such as lexicon, syntax, and complementation, and being involved in the conversation – has been studied. Specifically, vocabulary size (VS) was found to partly explain preschoolers' performances in False Beliefs tasks. While language fosters ToM development by putting one's own mental states and that of others into words, developing ToM abilities permits to be included more often in communication exchanges and also to be more effective while interacting with others (for a literature review, see Belacchi, 2022; Bulgarelli, Henning, & Bertin, 2022; Farrar et al., 2017).

Concerning gender differences in preschooler's ToM performances, a few studies found that girls were advantaged in False Belief tasks compared to boys, while often such differences were not detected (see Blijd-Hoogewys et al. 2022, for a review). Yet Charman et al. (2002) discussed that most of the studies lacked statistical power to demonstrate small gender differences.

Emotion Regulation, language, and gender

ER refers to the processes used to modify and manage the way emotions are experienced and how they are expressed through behaviour. The three core features of ER are the activation of a regulatory goal; the engagement of regulatory processes, that range from explicit, conscious, effortful, and

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

controlled regulation to implicit, unconscious, effortless, and automatic regulation; and the modulation of the emotion trajectory (Gross, 2024). In 1998, Gross elaborated a process model that highlights five families of emotional regulation processes in adults: situation selection to avoid unpleasant circumstances; situation modification, which are behavioural strategies such as physically removing from view the stimulus generating the emotion; attentional deployment, as searching for distraction to move the attention away from the situation; cognitive change to modify the emotional significance of a situation (which also includes cognitive reappraisal); and modulation of the emotional response itself.

ER strategies that are connected to Gross' model have been also used to describe ER in childhood. Attentional deployment seems quite stable in the preschool years, the use of behavioural strategies to regulate emotions improves from 3 to 6 years of age, and cognitive reappraisal appears at 3 years, if adults instruct the children to do so, and is spontaneously used from 5 years onwards (Davis et al., 2010; Sala et al., 2014).

In childhood, ER is measured through different methodological approaches: adult-report and self-report questionnaires, task-based procedures which often are designed to cause frustration or disappointment in the child, and narrative story stem tools (for a brief review, see Sala et al., 2014). Specifically, story stem completion procedures elicit children's narratives, to identify their representations through the analysis of both content and narrative style. The Emotion Regulation Story Stems (*Storie Interrotte per la Regolazione Emotiva*-SIRE; Sala et al., 2014) is one of these instruments. It allows to analyse and categorise the preschoolers' narratives about the regulation of five negative emotions, namely fear, anger, sadness, shame, and guilt. A detailed description of the test is reported in the Supplementary materials.

Language and ER are related: Eisenberg and colleagues (2005) proposed that linguistic skills promote ER, as children with higher communication abilities have more opportunities to learn about mental states, as just discussed. This model has been confirmed in subsequent studies: a small correlation between ER and receptive VS is observed in preschoolers (Cohen & Mendez, 2009; Reylly & Downer, 2019) and better language skills in toddlerhood predict the use of strategies to regulate emotion at 4 years of age (Roben et al., 2013). Cole et al. (2009) also found that expressive language predicts the generation of appropriate and effective strategies to regulate negative emotions (i.e., anger) in 3- and 4-year-olds. Ren et al. (2009) confirmed that both measures of expressive and receptive language are correlated to positive emotion regulation and emotion dysregulation in 3- to 5-year-olds.

Chaplin's bio-psycho-social contextual model is meant to explain the role

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

of gender on emotion expression: «Small biologically related gender differences in behaviours are present at birth, such as boys' greater reactivity and energy level and girls' greater language skills, and these elicit and are affected by socialization pressures from adults (and peers) to channel these differences into gender-role-consistent gender differences in emotion expressions as youth develop from infancy into the toddler/preschool period and childhood. In addition, there are contextual forces at work, shaping when and whether children express emotions according to gender roles depending on an interaction between their (biologically and/or socialization-influenced) tendency to express gender-role-consistent emotions and the particular social environment and larger cultural context in which they are found» (Chaplin, 2015, p. 10; see also Chaplin & Aldao, 2013). The model was derived from the meta-analysis of 164 studies, involving 21,709 participants from early infancy to adolescence. Girls display more positive emotions and internalizing emotions (e.g., sadness, anxiety, sympathy) than boys, and boys display more externalizing emotions than girls. Gender differences in positive emotions and externalizing emotions increase in middle childhood and adolescence. Children also use different display rules according to the person they are interacting with, and gender differences are less pronounced when children interact with parents, are more pronounced with unfamiliar adults (for positive emotions), and with peers or when they are alone (for externalizing emotions). In the literature, the effect of gender on ER development has not be detected systematically. In some studies, boys and girls did not show differences in their ER (Gilpin et al. 2015; Jahromi & Stifter, 2008), while in other research differences were found (Sala et al., 2014) in the use of ER strategies, with boys using more often behavioural strategies and girls referring more often to social support. Charman et al. (2002) discussion about the lack of statistical power to demonstrate small gender differences could also be the case for the research about ER in preschoolers.

Theory of Mind and Emotion Regulation

To our knowledge, only five studies investigated the relationship between ToM and ER in typically developing preschoolers: Baurain & Nader-Grosbois (2013), Gilpin et al. (2015), Hudson & Jacques (2014), Jahromi & Stifter (2008), and Lieberman et al. (2007). They involved in average 80 subjects (N range = 45-107) from 3 to 7 years of age; the participants were English-speaking (4 studies) or French-speaking children (Baurain & Nader-Grosbois, 2013). Three studies detected small to medium correlations between ToM and ER, where False Belief or ToM-emotion tasks were used to measure ToM, while ER scores were obtained during disappointment tasks

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

or dyadic games inducing positive versus negative emotions. According to Baurain & Nader-Grosbois (2013), children's performances on ToM tasks regarding causes and consequences of emotion correlated to two ER aspects coded by trained observers, namely emotional expression (r = .54) and behaviour towards social rules (r = .40). Hudson & Jacques (2014) found a small correlation between preschoolers' scores on a False Belief task and ER measured through disappointing gift task (r = .19). Lieberman et al. (2007) found that the association between False Belief tasks and ER, measured through a disappointment procedure and parents' report about their children emotional control, approached significance (r = .26, p < .054) only when controlling for verbal ability, measured with the PPVT-3 or PPVT-4. Two studies found no significant association between ToM and ER (Gilpin et al. 2015; Jahromi & Stifter, 2008): also in this case, False Belief tasks were used to assess ToM, and a disappointment task or a teacher-report tool was used for measuring ER.

Three of these studies also investigated the role of language on the relationship between ToM and ER, using different versions of the PPVT test, that measures the receptive VS of children: Gilpin et al. (2015) only detected an association between ToM and language (r = .43); Jahromi & Stifter (2008) and Lieberman et al. (2007) used receptive VS as a control variable in correlations between ToM and ER, and the first study found no association between the two constructs, while the second one found a positive correlation, as discussed above.

Of the five studies considered, two did not analyse gender differences in ToM or ER performance (Baurain & Nader-Grosbois, 2013; Lieberman et al., 2007), while three found no differences (Gilpin et al., 2015; Hudson & Jacques, 2014; Jahromi & Stifter, 2008).

The current study

Summarizing, ToM and ER Cognitive Reappraisal strategies develop from 3 to 6 years of age; both these competencies are related to language development and partly to gender, thus it is worth analysing the relationship between them. The present study aims to investigate the relationship between ToM and ER in typically developing preschoolers, also considering the role of age, receptive VS, and gender as control variables.

Our study might contribute to the current literature for many reasons. First, the scarce literature devoted to this topic shows mixed results, with some studies reporting positive correlations between ToM and ER, and others detecting a lack of association: our research might allow for a better understanding of the relationship between these constructs. Second, in this

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

paper ToM is assessed through a comprehensive test, the ToM Storybooks, consisting of repeated different tasks that allow a more reliable measure, as previously discussed. Third, ER is assessed with a narrative story stem procedure, that allows to directly observe and categorise the children's representations of ER strategies. Fourth, given the existing association between language and ToM or ER, and given that most of studies has been performed with English-speaking participants, it is worth testing the literature findings on different linguistic communities, as we did in this research conducted with Italian children.

Concerning the hypotheses of the current study, first, we expect an overall effect of age on ToM and language (specifically, VS), on ER *Behavioural* and *Cognitive Reappraisal Strategies* but not on ER *Attentional Deployment* and *Social Support* strategies. Second, we expect limited gender differences, particularly in ER *Attentional Deployment* and *Social Support* strategies, like in previous studies. Third, according to the existing literature, we expect language to be related to both ToM and ER. Fourth, given that the few studies in the literature analyzing the relationship between ToM and ER showed mixed results, we will explore the associations between the three ToM scores (*Quantitative*, *Qualitative*, and *Total*) and the four ER strategies.

Method

Participants

The participants were 62 Italian children (32 girls, 51.6%), aged from 3 to 6 years (range: 37-75 months, $M = 59.09$, $SD = 10.30$). Their average nonverbal IQ was 98.97 ($SD = 13.98$).

The inclusion criteria consisted of having a chronological age between 3 and 6 years and being enrolled in kindergarten; the main exclusion criterion was the presence of developmental difficulties or disabilities. Parents and teachers reported that the children had no health or mental issues. All the children were attending kindergarten and none of them had been already attending primary school. The participants of the current study were also included in the study by [Blind reference 1] and were part of the ToM Storybooks Italian standardisation sample [Blind reference 2].

Procedures and Measures

To take part in the study, the parents gave their written informed consent and the children their verbal consent. The research was conducted according

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

to the ethical principles of the Declaration of Helsinki. Data collection took place in 2011-2012. For the familiarisation, the administers of the tests spent a few hours with the children in the class to present themselves and freely play together. In the following weeks, the ToM Storybooks, the SIRE, and the PPVT-R were individually administered in the same period, in a quiet room at the kindergarten; the average time interval between the administration of the ToM Storybooks and the SIRE was 8 days (*SD* = 24 days).

The ToM Storybooks is a comprehensive test evaluating five ToM components: recognizing and naming emotions, making a distinction between physical and mental phenomena, understanding that perception leads to knowledge, understanding that desires and beliefs affect behaviours; a classic content False Belief task is also included. The test consists of six colour illustrated books telling the stories of Sam, a 5-year-old boy. Some tasks are proposed several times, to allow a more reliable measure. The administration requires about 40 minutes. The test has 95 items, that contribute to three scores. The *Quantitative score* varies from 0 to 77 and sums 77 close-ended questions. The *Qualitative score* varies from 0 to 36 and sums 18 open-ended questions. The children get 2 points if they spontaneously attribute mental states to Sam; 1 point if they simply explain situational aspects, or 0 points if they give incorrect answers. The *Total score* is the sum of the *Quantitative* and the *Qualitative score*. The ToM Storybooks have good internal consistency, test-retest reliability, inter-rater reliability, divergent and convergent validity (Blijd-Hoogewys et al., 2008; Bulgarelli et al., 2015; Bulgarelli, Testa, & Molina, 2022).

The Emotion Regulation Story Stems (*Storie Interrotte per la Regolazione Emotiva*-SIRE) is an experimental story stem procedure; cloth and wooden puppets are used to animate the stories (Sala et al., 2014). The order in which the narratives are told is always the same and follows the increasing complexity of the emotions cited in the different stories (fear, anger, sadness, shame, and guilt). The first story is a *warm-up story*: a narrative that has no implications with the theme of emotions. The administration of the procedure required about 20 minutes. The procedure was videotaped, transcribed, and coded by two independent judges, and the interrater reliability based on 20% of the tapes, evaluated through Cohen's K, ranged from .64 and 1.0 (Sala et al., 2014). The following coding system, developed by the Modal Model proposed by Gross (Gross, 2024), was used to assess the four ER strategies addressed in the children's responses: *Behavioural Strategy* was coded when the child refers to the protagonist's actions that are directed to change the situation in order to manage the emotion (e.g.: "He fixes the broken toy"); *Social Support* was coded when the narrative mentions the intervention of another character helping the protagonist to overcome the negative emotion

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

(e.g.: “The teacher comes and sits close to her”); *Attentional Deployment* was coded when the child responds refereeing to actions or thoughts distracting the protagonist and also the core of the narration from the emotion (e.g.: “He starts to play with the toys that he likes”); while Cognitive Reappraisal was coded when the child refers to protagonist’s thoughts or evaluations that, through the meaning’s modification of the situation, permit the regulation of the emotion (e.g.: “She understands that the next year she also will go to another school and she will find her friend again”). The coding system was meant to be mutually exclusive, and only one strategy was assigned to each of the child’s five responses. Thus, each strategy score varied from 0 to 5, because the child could have potentially used the same strategy in every answer.

The receptive VS was measured with the Italian version of the PPVT-R (Stella et al., 2000). Refer to the Supplementary Materials for detailed information about the tests.

Data analysis

We conducted a cross-sectional study, comparing different age groups. The raw scores of the three ToM scores, the four ER scores, and the VS score were used. Two different methods for age calculation have been used depending on the analysis. A categorical measure was used in ANOVA analysis. The children were divided into four age groups: 3-, 4-, 5-, and 6-year-olds; in the 3-year-old group, the children’s age ranged from 2 years, 6 months, and 1 day to 3 years and six months; the same strategy was followed for the other age groups. In the correlation analysis, we utilized an approximate decimal age, calculated from the date of birth to the date of measurement of the ToM Storybooks. First, we tested the presence of age and gender differences in the participants’ performances. One-way Anova, with Bonferroni post-hoc test, was run on ratio measures, i.e., the ToM and VS average scores, and non-parametric Kruskal-Wallis (for k-independent samples) or Mann-Whitney (for 2 independent samples) exact tests (Monte Carlo method) were conducted on the ordinal measures, i.e., the ER median scores. Pearson correlation was used to test the associations between the ratio measures, i.e., ToM scores, decimal age, and the other variables, while Spearman correlation was used for the associations between the ordinal measures, i.e., ER scores, decimal age, and other variables. Finally, partial correlations were run to test the associations between ToM and ER scores, controlling for age, and VS. The statistical analyses were conducted using SPSS 28.0 for Windows (SPSS Inc., Chicago, IL, USA).

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Results

Age differences in Theory of Mind, Emotion regulation, and Vocabulary Size

The descriptives of ToM, ER, and VS scores by age groups are reported in Table 1. The one-way Anova analysis showed that the ToM performances grew regularly as children got older: *Quantitative* score ($F(3,58) = 19.668$, $p < .001$), *Qualitative* score ($F(3,58) = 20.490$, $p < .001$), and *Total* score ($F(3,58) = 22.824$, $p < .001$). Among the ER scores, the Kruskal-Wallis exact test (Monte Carlo method) confirmed that the use of *Behavioural Strategy* improved with age ($H = 9.235$, $p = .023$), as well as the level of appeal to *Cognitive Reappraisal* strategies ($H = 11.261$, $p = .009$); both *Social Support* and *Attentional Deployment* did not grow as a function of age (all $p > .20$). Also VS significantly increased with age (one-way Anova, $F(3,58) = 7.660$, $p < .001$). Post-hoc comparisons between the age groups' scores are reported in Table 1 (see Tab.1).

The correlations between ToM, ER, VS scores, and age are reported in Table 2. The three ToM scores were strongly and positively associated with age; among the four ER scores, only *Behavioural Strategy* and *Cognitive Reappraisal* were positively and significantly associated with age. VS score also significantly varied as a function of age (see Tab. 2).

Tab. 1 – *Descriptives and age differences between the age groups' average scores (SD) of Theory of Mind, Emotion Recognition, and Vocabulary Size*

	Age groups			
	3 years (n = 6)	*4 years (n = 14)*	*5 years (n = 26)*	*6 years (n = 16)*
ToM Quantitative °	32.50 (6.31) [a,b]	42.57 (9.57) [c]	54.26 (6.35) [a,c]	56.56 (9.35) [b,c]
ToM Qualitative °	.33 (.52) [a,b]	3.00 (2.80) [c,d]	8.00 (3.63) [a,c]	12.63 (5.92) [b,d]
ToM Total °	32.83 (6.56) [a,b]	45.57 (12.16) [c,d]	62.26 (8.95) [a,c]	69.18 (14.13) [b,d]
ER Behavioural Strategy ^	.00 (.00) [a,b,c]	.57 (.65) [a]	1.15 (1.32) [b]	1.38 (1.41) [c]
ER Social Support ^	1.33 (1.21)	1.93 (1.64)	1.88 (1.42)	1.94 (1.12)
ER Attentional Deployment ^	.00 (.00)	.00 (.00)	.27 (.60)	.19 (.40)
ER Cognitive Reappraisal ^	.00 (.00)	.000 (.00) [a]	.27 (.60) [b]	.81 (1.05) [a,b]

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

VS°	23.50 (16.91) [a,b]	46.14 (18.97) [c]	56.38 (21.02) [a]	69.88 (23.35) [b,c]

Note: Within each score reported in the raw, corresponding superscript letters refer to significantly different groups (°Anova, Bonferroni test, $p < .05$; ^Mann-Whitney exact test, Monte Carlo method, $p < .05$).

Gender Differences in Theory of Mind, Emotion Regulation, and Vocabulary Size

No significant gender differences were found in the three ToM scores (One-way Anova; all $p > .60$) and in VS (One-way Anova; $p > .30$). As to the four ER scores, boys and girls differed in the use of *Behavioural Strategy*, as they appeared more frequently in the males' answers (boys: median = 1.00, girls: median = 0.00; Mann-Whitney exact test, Monte Carlo method, $U = 328.000$, $p = .022$). They also differed in the use of *Social Support*, that appeared more frequently in the females' answers (boys: median = 1.00, girls: median = 2.00; Mann-Whitney exact test, Monte Carlo method, $U = 327.000$, $p = .026$). No gender differences were found in *Attentional Deployment* ($p > .50$) and Cognitive Reappraisal ($p > .50$).

Tab. 2 – *Correlations between age, Vocabulary Size (VS), Theory of Mind (ToM), and Emotion Recognition (ER) scores*

		Age	**VS**	**ER BS**	**ER SS**	**ER AD**	**ER CR**
Correlations	*Age*	1.00	.565**	.370**	.085	.085	.462**
	VS	.565**	1.00	.423**	.119	-.129	.325*
	ToM Quant.	.727**	.684**	.310*	.104	.111	.409**
	ToM Qual.	.727**	.663**	.300*	.065	.120	.531**
	ToM Total	.763**	.711**	.322*	.095	.120	.473**
Partial correlations controlling for age, gender, and VS	*ToM Quant.*	--	--	-.035	-.052	.188	.108
	ToM Qual.	--	--	-.049	-.105	.191	.345**
	ToM Total	--	--	-.046	-.082	.219	.221

Note: BS = Behavioural strategy; SS = Social support; AD = Attentional deployment; CR = Cognitive reappraisal. Peason's correlations were run for Theory of Mind scores and Vocabulary Size score and Spearman correlations for Emotion Regulation scores; * $p < .05$, ** $p < .01$.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Relationship between Theory of Mind, Emotion Regulation, and Vocabulary Size

Table 2 also reports the correlations between the children's scores. The three ToM scores were significantly correlated to the ER *Behavioural Strategy* and the ER *Cognitive Reappraisal* scores; the second type of association was stronger. VS was significantly correlated to the three ToM scores and to ER *Behavioural Strategy* and ER *Cognitive Reappraisal* (see Table 2).

To better describe the relationship between ToM and ER scores, a partial correlation was conducted, controlling for age in months, IQ, and VS: only the correlation between the ToM *Qualitative* score and the ER *Cognitive Reappraisal* score was still moderate and significant (see Table 2).

Discussion

The current study investigated the relationship between ToM and ER, controlling for age, gender, and VS in 3- to 6-year-old Italian children. Our first hypothesis, based on the consolidated literature findings, was that the children's performances on different tasks would improve with age. The ToM Storybooks scores (*Qualitative*, *Quantitative*, and *Total*) and the VS score grew regularly with age, as expected from previous research using the same instruments (Blijd-Hoogewys et al., 2008; Bulgarelli et al., 2015; Lieberman et al., 2007). Also, the SIRE *Behavioural Strategy* and *Cognitive Reappraisal* scores were associated with age, showing that these two strategies are susceptible to age-related developmental processes in preschool years. Besides, the SIRE *Attentional Deployment* and *Social Support* scores were not correlated with age: probably, these two strategies depended more on socialization process (see below) or were more susceptible to inter-individual differences than to age. This pattern of results was also expected, as the current study involved the same children of [Blind reference 1]. Moreover, the literature devoted to ER development in preschool age systematically reported that this competence improved in the first years of life (Baurain & Nader-Grosbois, 2013; Lieberman et al., 2007).

The second hypothesis was also confirmed. No significant gender differences were found in the three ToM scores and in VS, coherently with previous research (Blijd-Hoogewys et al., 2022; Bulgarelli et al., 2015; Charman et al., 2002). As to the four ER scores, as expected, we observed the same pattern of results of Sala et al. (2014): the use of *Behavioural Strategy* was

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

more frequent in the males' answers, and the reference to *Social Support* appeared more frequently in the females' responses, while no differences were found in *Attentional Deployment* ($p > .50$) and *Cognitive Reappraisal*. Interestingly, ER *Social Support* strategy was susceptible to gender and did not grow with age, showing that, probably, socialisation processes played a crucial role in this case, as expected by Chaplin's bio-psycho-social contextual model of emotion expression (Chaplin & Aldao, 2013). The model was derived from the meta-analysis of 164 studies, involving 21,709 participants from early infancy to adolescence. Girls display more positive emotions and internalizing emotions (e.g., sadness, anxiety, sympathy) than boys, and boys display more externalizing emotions than girls. Children also use different display rules according to the person they are interacting with.

Considering the third hypothesis regarding language, we found that VS was significantly associated with the three ToM scores, as expected from the literature (Bulgarelli, Henning & Bertin, 2022; Milligan et al., 2007). VS was also associated with ER scores – namely *Behavioural Strategy* and *Cognitive Reappraisal* –, as already detected in previous studies (Cohen & Mendez, 2009; Gilpin et al., 2015; Ren et al., 2009; Reylly & Downer, 2019; Roben et al., 2013).

Moving towards the fourth main objective of the current study, our aim was to explore the interrelations between the ToM and the ER scores, as some previous findings showed that these constructs are correlated (Baurain & Nader-Grosbois, 2013; Hudson & Jacques, 2014; Lieberman et al., 2007), while other did not find such an association (Gilpin et al., 2015; Jahromi & Stifter, 2008). The three ToM scores (*Qualitative*, *Quantitative*, and *Total* scores) and the ER *Behavioural Strategy* and ER *Cognitive Reappraisal* scores were associated; actually, these scores also showed a clear developmental trend, as they were correlated with age. Conversely, the three ToM scores did not correlate with ER *Social Support* and *Attentional Deployment*, which were not associated with age and depended on gender.

Taking into consideration other variables that could influence the relationship between ToM and ER, we conducted partial correlations controlling for age, gender, and language: the correlation between the ToM *Qualitative* score and the ER *Cognitive Reappraisal* score turned out to be the only significant one. Also Lieberman et al. (2007) found a positive correlation between ToM and ER to approach significance ($r = .26$, $p < .054$) only when controlling for verbal ability, measured with the same tool that we used, the PPTV. Such a result is highly interesting, given the nature of the measures that we used in this study. In fact, the *Qualitative* ToM score depended on the ability of the child to reason about mental states: highest *Qualitative* ToM scores were reached when the child spontaneously talked about emotions,

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

desires, and beliefs, attributing them to the protagonists of the stories while answering to the open-ended questions of the ToM Storybooks. Coherently, the ER *Cognitive Reappraisal* score was attributed when the child referred to the use of cognitive abilities (thoughts, evaluations, etc.) that act on the meaning of the situation. Thus, ToM abilities measured through the ToM Storybooks were associated with more complex use of ER strategies in the SIRE story stem procedure.

The current study adds to the literature thanks to a series of methodological features. First, we used two new tools. The ToM Storybooks is a comprehensive test that consists of several tasks tapping five components of basic ToM. This characteristic allows for a more stable and reliable measure, that differentiates our study from the other five that already analysed the relationship between ToM and ER, using only False Belief or ToM-emotion tasks. Second, ER was assessed here with the SIRE story stem, and this makes the current study the first one to analyse ToM and ER association using a narrative procedure, which permits to observe how children develop the script and which kind of strategies to regulate emotion they use. In previous research, disappointment tasks, teacher-report tools, or dyadic games inducing positive versus negative emotions were used. Third, to our knowledge, this was the first study about the relationship among ToM, ER, and language to be conducted with Italian speaking children. As discussed in the introduction, ToM and language are inter-independent competencies (Milligan et al., 2007) and ToM processing during False Belief tasks was proved to partly depend on cultural and linguistic features, both when comparisons were made between Western and Eastern countries, and among Western European countries belonging to two cultural Latin and Germanic clusters (Blijd-Hoogewys et al., 2022; Wellman, 2018). Given that the literature devoted to the relationship between ToM, ER, and language is mainly run on English-speaking children, the current study contributes to filling a gap in the field.

Limitations of the current study

Concerning the limitations of our study, a rather small sample of children was involved (N = 62) and, specifically, the size of the younger group of participants could impact the robustness of the results. The study has an exploratory nature and the findings cannot be generalised to different populations.

The study was correlational, thus no casual relationship among the variables could be tested. For instance, language and ToM are inter-related competences, yet the impact of early language on future ToM abilities was proved to be stronger than the opposite (Milligam et al., 2007). Therefore,

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

future longitudinal research or intervention studies about ToM and ER could help better understand if ToM and ER are inter-related competence and if one of the two has greater impact on developing the other one. Also, future studies could further analyse the role of moderating variables on the relationship between ToM and ER, such as language and gender.

In summary, the current study showed that ToM and ER are associated competencies in preschool Italian children and that the child's ability to talk about emotions, desires, and beliefs spontaneously, is connected to higher cognitive ER strategies, even when language ability, gender, and age are controlled for.

Acknowledgments

We are grateful to the parents and children who participated in this study and to the teachers and the directors of the schools where the children were recruited. We also thank the students who contributed to the data collection.

Disclosure statement

No potential conflict of interest was reported by the authors.

Author contributions

Authors NS and PM designed the study. Author DB and NS collected the data. BR, DB, NS, and PM analysed the data. Author DB wrote the first versions of the paper and author BR, NS, and PM edited the final paper. All authors approved the final version of the manuscript for submission.

Data availability

Anonymised data are available from the authors upon reasonable request. Enquiries can be sent to the corresponding author.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

References

Baurain, C., & Nader-Grosbois, N. (2013). Theory of mind, socio-emotional problem-solving, socio-emotional regulation in children with intellectual disability and in typically developing children. *Journal of Autism and Developmental Disorders*, *43*, 1080-1097. DOI: 10.1007/s10803-012-1651-4.

Beaudoin, C., Leblanc, É., Gagner, C., & Beauchamp, M. H. (2020). Systematic Review and Inventory of Theory of Mind Measures for Young Children. *Frontiers in Psychology*, *10*(2905), 1-23. DOI: 10.3389/fpsyg.2019.02905.

Belacchi, C. (2022). Teoria della Mente e linguaggio: come condividere emozioni e pensieri con le parole [Theory of Mind and language: how to share emotions and thoughts with words]. In I. Castelli, D. Massaro, & A. Marchetti (Eds.), *La competenza sociale nei bambini* (pp. 101-134). Bergamo: Edizioni Junior-Bambini srl.

Blijd-Hoogewys, E. M. A., Bulgarelli, D., Molina, P., & van Geert, P. L. C. (2022). Nonlinearities in theory-of-mind development: New evidence from Dutch and Italian boys and girls. *European Journal of Developmental Psychology*, *19*(6), 905-924. DOI: 10.1080/17405629.2021.2020636.

Blijd-Hoogewys, E. M. A., van Geert, P. L. C., Serra, M., & Minderaa, R. B. (2008). Measuring theory of mind in children. Psychometric properties of the ToM Storybooks. *Journal of Autism and Developmental Disorders*, *38*(10), 1907-1930. DOI: 10.1007/s10803-008-0585-3.

Bulgarelli, D., Henning, A., & Bertin, E. (2022). Social cognition and its main correlates in childhood. In P. K. Smith & C. H. Craing (Eds.), *The Wiley-Blackwell Handbook of Childhood Social Development* (pp. 613-630). DOI: 10.1002/9781119679028.ch33.

Bulgarelli, D., Testa, S., & Molina, P. (2015). Factorial structure of the 'ToM Storybooks': A test evaluating multiple components of theory of mind. *British Journal of Developmental Psychology*, *33*(2), 187-202. DOI: 10.1111/bjdp.12062.

Bulgarelli, D., Testa, S., & Molina, P. (2022). Theory of mind development in Italian children with specific language impairment and autism spectrum disorder: delay, deficit, or neither?. *Journal of Autism and Developmental Disorders*, *52*(12), 5356-5366. DOI: 10.1007/s10803-021-05389-8.

Chaplin, T. M. (2015). Gender and emotion expression: A developmental contextual perspective. *Emotion Review*, *7*(1), 14-21. DOI: 10.1177/1754073914544408.

Chaplin, T. M., & Aldao, A. (2013). Gender differences in emotion expression in children: A meta-analytic review. *Psychological Bulletin*, *139*(4), 735-765. DOI : 10.1037/a0030737.

Charman, T., Ruffman, T., & Clements, W. (2002). Is there a gender difference in false belief development?. *Social development*, 11(1), 1-10. DOI: 10.1111/1467-9507.00183.

Cohen, J. S., & Mendez, J. L. (2009). Emotion regulation, language ability, and the stability of preschool children's peer play behavior. *Early Education and Development, 20*(6), 1016-1037. DOI: 10.1080/10409280903430745.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Cole, P. M., Dennis, T. A., Smith-Simon, K. E., & Cohen, L. H. (2009). Preschoolers' emotion regulation strategy understanding: Relations with emotion socialization and child self-regulation. *Social Development*, *18*(2), 324-352. DOI: 10.1111/j.1467-9507.2008.00503.x.

Davis, E. L., Levine, L. J., Lench, H. C., & Quas, J. A. (2010). Metacognitive emotion regulation: children's awareness that changing thoughts and goals can alleviate negative emotions. *Emotion, 10*(4), 498. DOI: 10.1037/a0018428.

Eisenberg, N., Sadovsky, A., & Spinrad, T. L. (2005). Associations of emotion-related regulation with language skills, emotion knowledge, and academic outcomes. *New Directions for Child and Adolescent Development*, *109*, 109-118. DOI: 10.1002/cd.143.

Farrar, M. J., Benigno, J. P., Tompkins, V., & Gage, N. A. (2017). Are there different pathways to explicit false belief understanding? General language and complementation in typical and atypical children. *Cognitive Development, 43,* 49-66.

Gilpin, A. T., Brown, M. M., & Pierucci, J. M. (2015). Relations between fantasy orientation and emotion regulation in preschool. *Early Education and Development*, *26*(7), 920-932. DOI: 10.1080/10409289.2015.1000716.

Gross, J. J. (2024). Conceptual foundations of Emotion regulation. In J. J. Gross & B. Q. Ford (Eds.), *Handbook of emotion regulation. Third edition* (pp. 3-12). New York, NY: Guilford Press.

Hudson, A., & Jacques, S. (2014). Put on a happy face! Inhibitory control and socioemotional knowledge predict emotion regulation in 5-to 7-year-olds. *Journal of Experimental Child Psychology*, *123*, 36-52. DOI: 10.1016/j.jecp.2014.01.012.

Jahromi, L. B., & Stifter, C. A. (2008). Individual differences in preschoolers' self-regulation and theory of mind. *Merrill-Palmer Quarterly (1982), 54*(1), 125-150. DOI: 10.1353/mpq.2008.0007.

Liebermann, D., Giesbrecht, G. F., & Müller, U. (2007). Cognitive and emotional aspects of self-regulation in preschoolers. *Cognitive Development*, *22*(4), 511-529. DOI: 10.1016/j.cogdev.2007.08.005.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*(2), 622-646. DOI: 10.1111/j.1467-8624.2007.01018.x.

Reilly, S. E., & Downer, J. T. (2019). Roles of executive functioning and language in developing low-income preschoolers' behavior and emotion regulation. *Early Childhood Research Quarterly*, *49*, 229-240. DOI: 10.1016/j.ecresq.2019.07.006.

Ren, Y., Wyver, S., Xu Rattanasone, N., & Demuth, K. (2016). Social competence and language skills in Mandarin-English bilingual preschoolers: The moderation effect of emotion regulation. *Early Education and Development*, *27*(3), 303-317. DOI: 10.1080/10409289.2015.1066639.

Roben, C. K., Cole, P. M., & Armstrong, L. M. (2013). Longitudinal relations among language skills, anger expression, and regulatory strategies in early childhood. *Child Development*, *84*(3), 891-905. DOI: 10.1111/cdev.12027.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Sala, M. N., Pons, F., & Molina, P. (2014). Emotion regulation strategies in preschool children. *British Journal of Developmental Psychology*, *32*(4), 440-453. DOI: 10.1111/bjdp.12055.

Stella, G., Pizzoli, C., & Tressoldi, P. E. (2000). *PPVT-R, Peabody Picture Vocabulary Test-Revised*. Torino: Omega Edizioni.

Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology*, *15*(6), 728-755. DOI: 10.1080/17405629.2018.1435413.

Supplementary materials

Procedures and Measures

Description of the tests

The ToM Storybooks is a comprehensive test evaluating five ToM components, based on Wellman's theory (1990): recognizing and naming emotions, making a distinction between physical and mental phenomena, understanding that perception leads to knowledge, understanding that desires and beliefs affect behaviors; a classic content False Belief task is also included. The test consists of six color illustrated books telling the stories of the protagonist Sam, a 5-year-old boy, his family, and friends. Each task is presented within the context of a story and some tasks are proposed several times, to allow a more reliable measure (Figure 1).

ToM Storybooks
Example of one task about
the role of belief on action

Next, Sam is going to feed the cows, together with uncle Bart.

"Can you get the bucket with feed?", uncle Bart asks.

The bucket with feed can be [point out] behind the tree or [point out] in the shed.

Sam wants to find the bucket.

He thinks the bucket is not behind the tree.

Q1. Where will Sam look?

Q2. Why is Sam looking …[there]?

Q3. Where does Sam think the bucket is?

Note: in this example, the open-ended answer to Q2 is scored 2 for "Because Sam ***thinks*** *the bucket is there (i.e., mental state)", 1 for "Because the bucket* ***is there*** *(i.e., factual situation)", 0 for "All the buckets are in sheds (i.e., general statement not taking into account Sam's thoughts or the story)" or if the child is incorrect at Q1.*

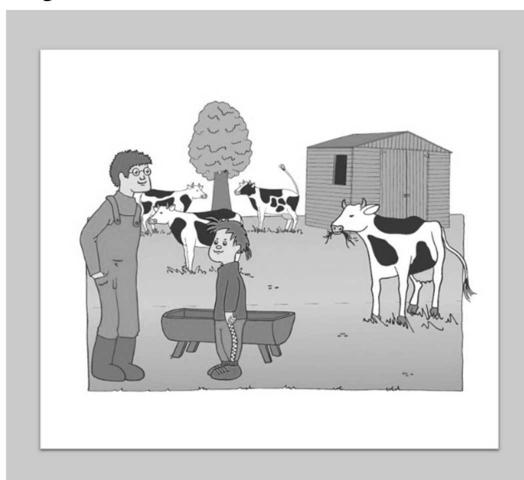

Figure 1. Example of one task of the ToM Storybooks

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

The administration of the test required about 40 minutes, inclusive of a short break after reading three books. The test has 95 items, that contribute to three scores. The Quantitative score varies from 0 to 77 and sums 77 close-ended questions, some of which require a non-verbal response (e.g., pointing images). The Qualitative score varies from 0 to 36 and sums 18 open-ended questions. As an example, to illustrate the scoring of these answers, let us take into consideration the False Belief task that is included in the test. The experimenter reads this story while the child is looking at the pictures: "Grandpa and grandma are paying Sam a visit. Sam gets rollerblades from grandpa and grandma. […] Sam puts the rollerblades [point out] in the toy trunk. Then, he goes upstairs to play with his crane. When Sam has left, his sister goes to the toy trunk. She likes to tease her brother. Lotje hides the rollerblades [point out] in the box! And then, she goes quickly outside. Then, Sam is coming back. He wants to roller skate. Where will Sam look for his rollerblades? Why is Sam looking [there]?". The children get 2 points if they spontaneously attribute mental states to Sam (e.g.: "Sam looks for the skates in the box because he thinks they are there"), 1 point if they simply explain situational aspects (e.g.: "Sam looks for the skates in the box because he put them there"), or 0 points if they give incorrect answers (e.g.: "Sam looks for the skates in the toy trunk because they are there"). The Total score is the sum of the Quantitative and the Qualitative score. The ToM Storybooks have been standardized in the Netherlands and Italy and it has good internal consistency, test-retest reliability, and inter-rater reliability (Blijd-Hoogewyset al., 2008; Bulgarelli et al., 2015). Both the Dutch and the Italian version show a five-component factorial structure, that are similar: for a discussion about the two factorial structures, see Bulgarelli et al. (2015). The test also has discriminant validity, as it differentiates between autistic and typically developing children, and divergent and convergent validity (Blijd-Hoogewys et al., 2008; Bulgarelli, Testa & Molina, 2022).

The Emotion Regulation Story Stems (*Storie Interrotte per la Regolazione Emotiva*-SIRE) is an experimental story stem procedure; cloth and wooden puppets are used to animate the stories (Sala et al., 2014). The procedure was derived from the story stem tools developed to research attachment representations (Bretherton et al., 1990; Green et al., 2000). One warm-up and five emotional stories take place in the school context; such an environment was chosen for two reasons: first, it permits to investigate the theme of ER without directly involving the representations of the attachment to parents; second, kindergarten was indicated as the first context in which children's ability to regulate emotions take place in addition to the family(Silvers et al., 2012). The order in which the narratives are told is always the same and follows the increasing complexity of the emotions cited in the different

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

stories (fear, anger, sadness, shame, and guilt). The protagonist of every tale is always the same character, that can be a boy or a girl corresponding to the gender of the child who answers to the procedure. Coherently with other story stems developed in the field of attachment research, such as ASCT (Bretherton et al., 1990), the first story is a warm-up story: a narrative that has no implications with the theme of emotions. The researcher initially tells the warm-up story for the purpose of establishing contact with the child and in order to give her/him the opportunity to familiarize with the task, the context, and the materials. The warm-up story describes an ordinary day at school in which children are going to play in the yard (Figure 2).

Figure 2. Warm-up story of the SIRE

Subsequently, the researcher tells five stories related to the five emotions cited. In the story associated with fear, the protagonist gets lost during school trip. In the narrative related to anger, the protagonist's companions make her/him drop her/his snack. The best friend of the protagonist is moving to another city and leaves the classroom in the story associated with sadness. The following narrative involves the child falling in front of people and feeling ashamed (Figure 3).While in the last story the protagonist breaks a friend's favorite toy and feels guilty. All the stories are narrated both verbally

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

and by moving the puppets simultaneously. At the end of each story, the researcher labels the emotion involved and asks the child to continue the narrative. The administration of the procedure required about 20 minutes, depending also on the child's playfulness, responses, and narrative skills. The procedure was videotaped, transcribed, and coded by two independent judges, and the interrater reliability based on 20% of the tapes, evaluated through Cohen's K, ranged from .64 and 1.0 (Sala et al., 2014).

Figure 3. Shame story of the SIRE

The following coding system, developed by the Modal Model proposed by Gross (Gross, 1998; 2014), was used to assess the four ER strategies addressed in the children's responses: Behavioral Strategy was coded when the child refers to the protagonist's actions that are directed to change the situation in order to manage the emotion (e.g.: "He fixes the broken toy"); Social Support was coded when the narrative mentions the intervention of another character helping the protagonist to overcome negative emotion (e.g.: "The teacher comes and sits closet to her"); Attentional Deployment was coded when the child responds refereeing to actions or thoughts distracting the protagonist and also the core of the narration from the emotion (e.g.: "He starts

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

to play with the toys that he likes"); while Cognitive Reappraisal was coded when the child refers to protagonist's thoughts or evaluations that, through the meaning modification of the situation, permit the regulation of the emotion (e.g.: "She understands that the next year she also will go to another school and she will find her friend again"). The coding system was meant to be mutually exclusive, and only one strategy was assigned to each of the child's five responses. Thus, each strategy score varied from 0 to 5, because the child could have potentially used the same strategy in every answer.

The receptive VS was measured with the Italian version of the PPVT-R (Dunn & Dunn, 1981; Stella et al., 2000), a well-known test for evaluating receptive language. It asked the child to recognize the image associated with a target word read by the interviewer; the image was inserted on a page together with three distractors. The tests consisted of 175 tables, of which target words were increasingly more difficult. Thanks to the evaluation of a floor line and a ceiling line, the administration of the test requires about 15 minutes.

References

Blijd-Hoogewys, E. M. A., Van Geert, P. L. C., Serra, M., & Minderaa, R. B. (2008). Measuring theory of mind in children. Psychometric properties of the ToM Storybooks. *Journal of Autism and Developmental Disorders, 38*(10), 1907-1930. DOI: 10.1007/s10803-008-0585-3.

Bretherton, I., Ridgeway, D., & Cassidy, J. (1990). Assessing internal working models of the attachment relationship: An attachment story completion task for 3-year-olds. In M. T. Greenberg, D. Cicchetti & E. M. Cummings (Eds.), *Attachment in the preschool years: Theory, research, and intervention* (pp. 273-308). Chicago, IL US: University of Chicago Press.

Bulgarelli, D., Henning, A., & Bertin, E. (2022). Social cognition and its maincorrelates in childhood. In P. K. Smith & C. H. Craing (Eds.), *The Wiley-Blackwell Handbook of Childhood Social Development* (pp. 613-630). DOI: 10.1002/9781119679028.ch33.

Bulgarelli, D., Testa, S., & Molina, P. (2015). Factorial structure of the 'ToM Storybooks': A test evaluating multiple components of theory of mind. *British Journal of Developmental Psychology, 33*(2), 187-202. DOI: 10.1111/bjdp.12062.

Dunn, L M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test Revised*. American Gudance Service.

Green, J., Stanley, C., Smith, V., & Goldwyn, R. (2000). A new method of evaluating attachment representations in the young school-age children: The Manchester Child Attachment Story Task (MCAST). *Attachment & Human Development, 2*(1), 48-70. DOI: 10.1080/146167300361318.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*, 271-299. DOI: 10.1037/1089-2680.2.3.271.

Gross, J. J. (2014). Emotion regulation: Conceptual and empirical foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 3-20). New York, NY: Guilford Press.

Sala, M. N., Pons, F., & Molina, P. (2014). Emotion regulation strategies in preschool children. *British Journal of Developmental Psychology, 32*(4), 440-453. DOI: 10.1111/bjdp.12055.

Silvers, J. A., McRae, K., Gabrieli, J. D. E., Gross, J. J., Remy, K. A., & Ochsner, K. N. (2012). Age-related differences in emotional reactivity, regulation, and rejection sensitivity in adolescence. *Emotion, 12*(6), 1235-1247. DOI: 10.1037/a0028297.

Stella, G., Pizzoli, C., & Tressoldi, P. E. (2000). PPVT-R, Peabody Picture Vocabulary Test-Revised. Torino: Omega Edizioni.

Wellman, H. M. (1990). *The child's theory of mind*. Cambrige, MA: MIT Press.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Teaching Clinical Reasoning and Critical Thinking in Psychology Graduate Courses: A Systematic Review

Tugce Elbir *, Catia Cammeo **, Francesca Cavallini *,
Antonella Marchetti °, Davide Massaro °

* TICE Cooperativa Sociale,
Piacenza, Italy;
e-mail: tuceelbir@gmail.com;
e-mail: francesca.cavallini01@gmail.com.
** University Magna Graecia of Catanzaro, Italy;
e-mail: catia.cammeo@gmail.com.
° Catholic University of the Sacred Hearth, Milan, Italy;
e-mail: antonella.marchetti@unicatt.it;
e-mail: davide.massaro@unicatt.it.

Ricevuto: 15.07.2021 - **Accettato:** 22.05.2024

Pubblicato online: 23.12.2024

Abstract

Introduction: There are many studies and reviews that investigate the most functional ways of explicitly teaching clinical reasoning skills in more traditionally healthcare areas such as medicine and nursing, hypothesizing that reducing the gap between theory and practice could guarantee continuity in the quality of care provided to patients.

Clinical reasoning and critical thinking are aspects that in clinical practice guide the psychologist in diagnosis and intervention. However, in the psychological field there is a lack of experimental studies that address the issue of direct or explicit teaching of these skills.

This systematic review study sought to investigate, via a systematic review, whether there is an association between the teaching methodology and the student's improvement in their critical analyses and clinical reasoning skills.

Statement of the Problem: Despite the importance for clinicians, the interest of researchers and academics is quite recent. Teaching of clinical reasoning is, on the contrary, more common in medicine and nursing sectors.

Elbir et al./ *Ricerche di Psicologia, 2024, Vol. 47*
ISSNe 1972-5620, Doi:10.3280/rip2024oa18274

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Methodology: Articles published from 2004 until 2023 within PubMed (National Center for Biotechnology Information, NCBI), Web of Science (Core Collection, Clarivate Analytics), Scopus (Elsevier), EBSCO, Scielo, and Psychinfo databases were searched through the following keywords: Critical Thinking - Clinical Reasoning; Problem Based Learning - Case Based Learning; Teaching Methods - Teaching Strategies; Learning Strategies - Clinical Learning; University; Intervention - Training; Psychology; Psychology Students.

Results: Theoretical and practical implications to teach clinical reasoning and critical thinking in the psychological field were discussed, including recommendations for designing researches aimed at investigating the impact of the different methodologies, as well as the possibility of including education in critical thinking within the different levels of training dedicated to psychologists.

Conclusion: The collected studies presented methodologies that, although different, significantly promoted the improvement of critical thinking in university students, demonstrating that it is "possible" to transfer this complex skill.

Keywords: clinical reasoning, critical thinking, clinical learning, psychology students, teaching methods

Clinical Reasoning and Critical Thinking in the Health Professions

The following review originates from the authors' interest to investigate the theme of teaching clinical reasoning and critical thinking in Psychology graduate courses. As a matter of fact, a great deal of the psychological know-how consists of immaterial aspects, such as clinical reasoning, critical thinking and judgement, problem solving which in the clinical practice, guide the psychologist in the choices of diagnostic and intervention (Norman, 2000). Despite the importance of such constructs in infrastructures and for clinicians who work in the territory, the interest of researchers and academics is quite recent (Andrews & Syeda, 2016). Teaching of clinical reasoning and other non-conceptual skills is, on the contrary, more common in areas traditionally dedicated to healthcare, such as the medicine and nursing sectors (Burbach et al., 2015; Norman et al., 2018; Pelaccia et al., 2020).

There is a large number of studies and reviews addressing the most functional methods of explicit teaching of clinical reasoning skills in the healthcare sector (Abrami et al., 2008; Delany et al., 2020; Mason et al., 2020; Pramila-Savukoski et al., 2019; Richmond et al., 2020). These

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

studies have substantially increased in the last few years, also due to the pressure exerted by the emerging needs of healthcare infrastructures: employee turnover, massive retirement and lack of formal tools of transition to the labour market impose that, after completion of studies, the practitioner should be able to not only do, but also "think" of what to do (Cavallini et al., 2019; Torre et al., 2020). This systematic review/survey is inspired by the work of Payan-Carreira et al. (2019), a systematic literature review which presents and discusses teaching methodologies and assessment tools for clinical reasoning and critical thinking in healthcare.

The objective is to develop a first formal and systematic reflection on the educational practices of clinical reasoning and critical thinking in psychology. Critical thinking will be defined as the skill set that lead to a process of reflection which translates into interpretation, analysis and evaluation of data (for instance, diagnostic and criterial tests) in order to reach a judgement (Payan-Carreira et al., 2019; Sternberg & Halpern, 2020).

In literature the terms clinical reasoning, clinical judgement, problem solving, decision-making process and critical thinking are often used interchangeably (Faucher, 2011; Payan-Carreira et al., 2019; Shin, 2019). They are defined as the mindset that clinicians get into when thinking about the issues they face in their practice. It deals with clinical judgements (making a decision regarding what is right and what is wrong) and clinical decision-making process (deciding what to do), (Guerrero, 2019; Payan-Carreira et al., 2019; Royce et al., 2019). To summarize, clinical reasoning and critical thinking emphasize slightly different aspects but are often considered synonyms; research and reviews author, however, are aware of the differences between the two constructs, at theoretical level as well as in the ways of measurement and evaluation (Faucher, 2011; Payan-Carreira et al., 2019). In order to clarify this topic, Joyce (2013) developed a systematic review focused especially on the different ways in which various authors define the processes of critical thinking and clinical reasoning, focusing on the specificities of each construct and presenting the reader with the diverse intervention and assessment tools.

Although it presents differences and specificities, the review highlights how in synthesis tasks it is possible to use the terms interchangeably. As it is the case of this survey itself, they have been used indistinguishably in research with the scope of including as many relevant publications as possible.

Clinical Reasoning is conceptualized as that process through which healthcare professionals generate clinical judgement by choosing from

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

the alternatives, weighing the evidence, using intuition and recognizing reference models. It is a logical process through which healthcare specialists gather clues, elaborate information, come to an understanding of a problem affecting a patient or a situation, plan and carry out interventions, assess the results and reflect and learn the process (Faucher, 2011; Koufidis et al., 2020; Payan-Carreira et al., 2019; Robertson, 2012). Thus, clinical reasoning results in an intellectual activity which belongs to any profession and that, starting from a variety of information, leads to the identification of problems that are objects of interest (Kosior et al., 2019; Young, Thomas, Gordon et al., 2019). It is a form of circumstantial thought where clues are observable elements that stimulate to further research for new information and is based on the hypothetical-deductive process through the generation of one or more hypotheses (Al Rumayyan et al., 2018; Huhn et al., 2019). In the field of psychology this process starting from a health issue, an attitude, behavior or thought, passing through the medical condition, anamnesis and individual history leads to the formulation of one or more diagnostic hypotheses.

Critical thinking and clinical reasoning have in common a number of factors: the intentional commitment to ask clear and well-formulated questions, the gathering and assessment of relevant information, the care in thinking of available alternatives, the ability to recognize and assess hypotheses by taking into account the implications and practical consequences, hence the communication with other professionals (Payan-Carreira et al., 2019; Rutter & Harrison, 2020).

In the last decades the acquisition of clinical reasoning and critical thinking competencies (particularly for doctors, dentists, nurses, neuro psychomotricists and speech therapists) is once again of interest to higher education institutions (Choi et al., 2020; Ihm et al., 2020; Mutter et al., 2020; Pierce et al., 2020; Robertson, 2012). This interest and studies are aimed at the analysis of teaching practices which can transfer useful skills to facilitate the transition of recent graduates to the labour market (Vidyarthi et al., 2016; Young, Thomas, Lubarsky et al., 2020).

In the healthcare sector reducing the gap between theory and practice could favor recent graduate's integration in the work-team and guarantee the continuity of quality of care provided to patients (Durning et al., 2019). While these considerations are valid for the medical field (doctors and nurses) and in all historically health professions (speech therapists, occupational therapists, physiotherapists), specifically in the psychological field we witness the great shortage of experimental studies tackling the theme of direct or explicit teaching of the abilities described.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Data concerning the increase of mental health issues related to the Covid 19 pandemic (Li et al., 2020) suggest that the request of psychologists on the territory will undergo an increase; therefore, it becomes relevant to explore the teaching practices and methodologies that promote the acquisition of fundamental skills in the clinical practice. This systematic revision intends to evaluate the status of current teaching practices in use and tested to promote the acquisition or the improvement of critical thinking and clinical reasoning in psychologists or psychologists in training.

Method

In order to assess the current state of research within the area of teaching and education with regards to clinical reasoning and critical thinking in the psychological sector, we have worked on a systematic review of the literature in Italian and English languages published from 2004 until 2023.

The research has been carried out within five online databases: PubMed (National Center for Biotechnology Information, NCBI), Web of Science (Core Collection, Clarivate Analytics), Scopus (Elsevier), EBSCO, Scielo (Scientific Electronic Library Online) and Psychinfo.

The following Boolean search phrases combination was used: (Critical Thinking "OR" Clinical Reasoning) AND (University "OR" Intervention "OR" Training) AND (Psychology); (Problem Based Learning "OR" Case Based Learning) AND (Critical thinking "OR" Clinical Reasoning) AND (Teaching Methods "OR" Teaching Strategies) AND (Psychology Students); (Critical Thinking "OR" Clinical Reasoning) AND (Learning Strategies "OR" Clinical Learning) AND (Psychology Students).

Inclusion criteria applied in this review were outlined before searching for literature, and developed by means of PICOS (Methley et al., 2014):

- P (Population) – Students enrolled in Graduate courses (Bachelor or specialized training), Master or PhD in psychology or Postgraduate traineeship;
- I (Intervention) – Educational strategies explicitly used to teach clinical reasoning and critical thinking;
- C (Comparison) – Assessment tools and educational strategies;
- O (Outcomes) – The effective development of skills based on different educational strategies and assessment tools;
- S (Study Design) – Any quantitative study.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Every article has been jointly analysed by two reviewers. A first analysis was performed on 91 research articles published between 2004 and 2023. Fifty-three studies were excluded from the analysis as books, book chapters, dissertations, reviews, non-empirical studies, articles that did not investigate psychology students but those of medicine and nursing (or other professions) and which did not examine students but professionals.

In the second analysis, 35 articles were screened, of which 25 excluded for skills and dispositions not evaluated (clinical reasoning and critical thinking), subject and field of study not relevant to the analysis conducted.

Nine articles were evaluated for eligibility and included in the quantitative summary (Figure 1).

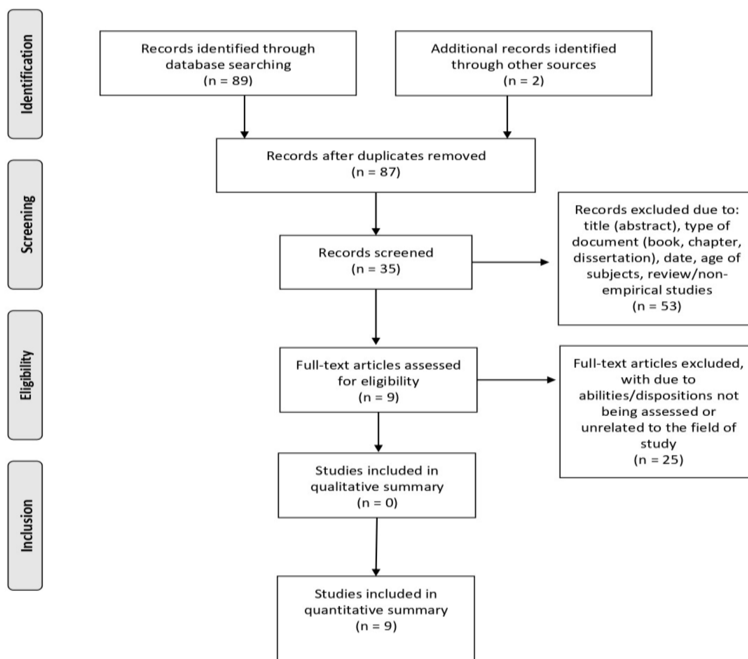

Fig. 1 - *Systematic review steps flowchart*

The extracted articles have been analysed by using a supplementary table (Table 1), created to recover relevant data.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Tab. 1 - *Supplementary table created to recover relevant data*

Authors	Participants	Assessment	Intervention	Results
Penningroth, S.L., Despain, L.H., & Gray, M.J. (2007)	47 students from the Psychological Sciences (PS) course group and 119 students from the comparison group enrolled in a General Psychology (GP) course	Psychological Critical Thinking Exam (PCTE), (Lawson, 1999)	Small team activities and group discussions, frequent close-ended questions and a final written task which required comparing the scientific research results to the evidence	Results show an increase in critical thinking skills
Solon, T. (2007)	51 Psychology students	Cornell (Z) Critical Thinking Test (Ennis et al., 1985) Textbooks and study guides exercises The 5 methods of Mill Socratic interrogation technique to stimulate critical thinking	The experimental group undergoes general critical thinking learning through a class laboratory and homework	The study demonstrates a significant and substantial growth of the students' critical thinking skills
Bensley, D.A., Crowe, D.S., Bernhardt, P., Buckner, C., Allman, A.L. (2010)	47 psychology students	Three Critical Thinking (CT) psychological test subscales (Bensley & Baxter, 2006) Version of 20 items of the reduced form of the Need for Cognition Scale (Cacioppo & Petty, 1982) Checklist for educational background assessment GPA to assess academic results SAT scores to assess academic skills	Students from the SRM class have analysed a literature review, completed homework and took quizzes on these abilities, and hence received corrective feedback about all their work. Students from traditional RM lectures have received instructions focused on learning statistics, design and research methodology, and on how to write a research report in the style of the American Psychological Association (APA)	Results support the efficacy of explicit teaching of critical thinking skills infused directly into normal course lectures
Haw,J. (2011)	84 second year students and 60 fourth year Psychology students	Psychological Critical Thinking Exam (PCTE), (Lawson, 1999)	Education-based teaching practice designed to improve psychological critical thinking skills	Results show a significant increase of critical thinking abilities and confirm the value of the education-based teaching practice

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Karantzas, G.C., Avery, R.M., Macfarlane, S., Mussap, A., Tooley, G., Hazelwood, Z., & Fitness, J. (2013)	273 third year psychology students	Self-efficacy Learning scale (Zimmerman, Kitsantas, & Campillo, 2005) Academic Motivation scale (AMS-C 28, Vallerand et al., 1992) Study process revised questionnaire (R-SPQ-2F, Biggs et al., 2001) Measure consisting of seven items specifically developed to study critical analysis and problem-solving skills	Collaborative learning approach (CLA) and use of a tutorial program created to encourage the development of university students' critical analyses and problem-solving skills. The program is inspired by the "Choose Your Own Adventure" (CYAO) series of novels	Students showed an improvement in their critical analyses and problem solving
Muehlenkamp, J.J., Weiss, N., & Hansen, M. (2015)	74 Psychology students enrolled in a Psychology introductory course	Five self-report questions from the interview Cognitive Subscale of the Student Engagement (Ahlfeldt et al., 2005) Five scenarios adapted from "Lawson's (1999) psychological critical thinking" Items from the Student Course Engagement Questionnaire (SCEQ), (Handelsman et al., 2005)	One of the authors, after examining various resources of PBL (es. Duch et al., 2001), has structured the 16 weeks Psychology course introduction into 4 learning units based on problems, and one learning unit generated by the students	Results show a significant increase of critical thinking skills, psychological research analysis abilities, and emotional involvement
Wentworth, D.K., Whitmarsh, L. (2017)	275 Psychology students, 89 of which have completed all three assessments	Questionnaire (of 13 items) developed to access students' reactions to each written task Instructor's Manual for Psychology, 12th Edition (Wade, Tavris & Garry) Scoring index created for studying (3 points scale)	Three innovative written homework have been developed in order to teach students to think like a psychologist, and they consist of: increasing critical thinking, apply research concepts and resist to plagiarism	Results show an increase in students' ability to critically think

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Campbell, C.G., & Oswald, B.R. (2018)	First course: 10 students (5 graduates and 5 undergraduates) Second course: 7 students (5 graduates and 2 undergraduates)	P-SAP Test (Fitch & Steinke, 2013)	Three series of activities and homework have been developed in order to maximize opportunities of students' engagement in reflections, and to provide a supporting structure for the use of critical thinking during such reflections	The study shows a significant improvement in students' critical thinking skills
Cammeo, C., Prestera, G., Massaro, D., Marchetti, D., Cavallini, F. (2022)	Total of 92 students enrolled in a Master Degree Course in Psychology of Clinical and Social Intervention at the University of Parma 73 of them attended lessons (attending group) 19 were non attending students (following the course only in telematic mode, and therefore not accessing the university classroom in person)	Pre and post-test 60 flashcard containing diagnostic criteria of neurodevelopmental disorders 40 tracks of the third proof (clinical case) of the State Examination for the qualification to the profession of psychologist Textbook 10 podcasts based on 10 different clinical cases	Attending group was exposed to clinical reasoning listening to the podcasts in classroom Non attending group used a more traditional methodology of studying trough the textbook Both the groups completed the clinical case proof	The findings revealed an improvement in the acquisition of clinical reasoning by the group of students exposed to podcast training in contrast to their colleagues who had studied written materials

Note: Studies included in systematic review

For the purpose of this review a large variety of instruments were used to measure clinical reasoning and critical thinking, and were classified as follows: General standardized tests (indicated in three articles); domain specific standardized tests (cited in six articles) – specifically used to measure abilities of and dispositions to clinical reasoning and critical thinking; domain specific indices, surveys or questionnaires (indicated in five articles) – usually not standardized, and of self-report type, created or used by researchers to specifically evaluate the experience of learning, that is the students' preservation of knowledge, self-confidence, satisfaction; podcast created ad hoc (cited in one article).

As for the improvement of abilities of or disposition to clinical reasoning, results have been classified as follows: general profit (++) – when a statistically significant profit in terms of a general skills set has been reported; specific profit (+) – when a statistically significant profit in terms of a specific skill has been reported; no profit (-) – when no

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

statistically significant profit has been mentioned, in terms of a specific ability, or a series of skills (clinical reasoning and critical thinking), or disposition.

Discussion

Population

The nine studies report a sample of 1474 total subjects, 833 of which are female and 321 are male subjects, with an age spanning between 18 and 25 (average age = 21.16; 6.08% 18 years old; 41.85% 19 years old; 0.90% 20 years old; 39.55% 22 years old; 4.19% 23 years old and 7.40% 25 years old). Unspecified were: the age and sex of a control group consisting of 119 subjects (Penningroth et al., 2007), the age of 47 subjects (Bensley et al., 2010) and the sex of 154 subjects who only participated to the initial phase of the study (Haw, 2011). All subjects attended a university psychology course. The groups varied in size; the most numerous consisted of 275 students and the less numerous consisted of seven students (M = 98.71; SD = 96.32). Among the total number of students who took part in the research, 583 were in their first year (47.94%), 362 in their second (22.20%), 273 in their third (22.45) and finally, 90 in the fourth year of the course (7.40%), (M = 304.0, SD = 204.73).

Types of intervention

All interventions involved the use of specific critical thinking teaching material (textbooks, workbooks, manuals, checklists, podcast) associated with various educational methodologies, such as: Problem Based Learning, Active Learning, Group discussion.

Four studies described non-specific methodologies which included guided discussions and group exercises, written as well as oral. One study included technological methodologies such as podcast (Cammeo et al., 2022).

Tools

In terms of the variables that were considered in this review of critical analysis, critical thinking and problem solving, different tools have been used: Standardized PCTE-Psychological Critical Thinking Exam-questionnaires (Haw, 2011; Muehlenkamp et al., 2015; Penningroth et al., 2007) in three studies; and CT tests (Bensley et al., 2010), Cornell Z (Solon, 2007) and P-SAP (Campbell & Oswald, 2018), each in one

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

study. The PCTE (Haw, 2011; Muehlenkamp et al., 2015; Penningroth et al., 2007) is the only tool specifically dedicated to critical thinking in the psychological area.

The other tools are also used in different areas. Ad hoc self-report questionnaires, in which the students had to assess their own critical thinking skills, in three studies (Muehlenkamp et al., 2015; Karantzas et al., 2013; Wentworth & Whitmarsh, 2007); furthermore, one study (Wentworth & Whitmarsh, 2007) utilized an ad hoc questionnaire to verify students' proficiency in identifying errors of thought. In two studies, different tools were used: in one study (Muehlenkamp et al., 2015), critical thinking has been investigated through students' self-evaluation as well as specific test; while in the study by Wentworth and Whitmarsh (2007), critical thinking has been investigated through self-evaluation and an ad hoc questionnaire.

Other variables under examination were: Students' motivation and participation and Learning. Students' motivation and participation have been investigated through standardized questionnaires (Bensley et al., 2010; Karantzas et al., 2013; Muehlenkamp et al., 2015); or through unstructured observation (Campbell & Oswald, 2018). In one study (Karantzas et al., 2013), learning style (CLA) was investigated, while in two studies (Bensley et al., 2010; Solon, 2007) learning of course-specific content was assessed, in addition to the general performance of the student investigated in the study of Bensley, et al. (2010).

Experimental designs

Two studies (Karantzas et al., 2013; Wentworth & Whitmarsh, 2017) have used an intra-group experimental design. Five studies are quasi-experiments where the element of random allocation, specifically, is missing. These are studies that analyze university students in different years of the course (Bensley et al., 2010; Campbell & Oswald, 2018; Haw, 2011; Penningroth et al., 2007; Solon, 2007). Finally, only two studies (Muehlenkamp et al., 2015; Cammeo et al., 2022), also structured as an intra-group study, presents a control group with random allocation.

Duration

Only one study did not specify the duration of interventions (Karantzas et al., 2013). The rest of the studies, on the other hand, can be separated in two categories: in six studies, the duration of interventions was expected to be around a semester (between 12 and 16 weeks); in the remaining two studies intervention duration would last one year (Muehlenkamp et al., 2015; Wentworth & Whitmarsh, 2017).

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Results obtained by researchers

Every article measured the statistical significance of the results obtained within the specific abilities as well as other dispositions (Bensley et al., 2010; Campbell & Oswald, 2018; Haw, 2011; Karantzas et al., 2013; Muehlenkamp et al., 2015; Penningroth et al., 2007; Solon, 2006; Wentworth & Whitmarsh, 2017; Cammeo et al., 2022).

Although the majority of studies has focused on detecting changes in the target ability (critical thinking, clinical reasoning), some authors have also investigated the effects of training on other abilities or looked for correlations between the target abilities and other variables.

The results obtained by Karantzas et al. (2013) show that an approach inspired by Problem Based Learning (Barrows, 1996), integrated with Collaborative Learning Approach (Dillenbourg, 1999; Roschelle & Teasley, 1995), promotes a significant change in the perception of one's ability to critically analyse and solve problems, as seen in 273 psychology students after just 10 weeks of intervention. The intervention involved students in structured exercises where, in small groups, they had to solve daily dilemmas and/or put themselves in a familiar psychologist's shoes. Moreover, thanks to the use of the Latent Growth Curve Modelling, the study (Karantzas et al., 2013) allowed to isolate the changes in critical analysis and problem-solving abilities before, during and after the treatment from typically connected variables, such as self-efficacy, intrinsic motivation and learning style. The authors highlight the exclusive use of self-reporting tools as the main limitation.

Professors often assign writing homework to students. The study by Wentworth and Whitmarsh (2017) has the objective of verifying the efficacy of three writing tasks in promoting the typical skills of psychological thought, which are: increasing critical thinking, applying research concepts and resisting plagiarism. The writing tasks required extremely personal reasoning and analysis, so they made it unlikely and unnecessary for the student to "copy". Critical thinking has been assessed through a self-report tool as well as questionnaires in which each student had to consider a certain scenario and highlight the errors in the ways of thinking. The outcome has shown that both measures of critical thinking improve significantly from the first to the second written homework and from the second to the third, while there are no improvements from the first to the second: the authors speculate that in order to obtain a change in the variables under consideration, further exercising is needed. Besides the positive results in critical thinking, the authors emphasize the almost absence (only one student out of 456) of plagiarism behaviour in the writing homework execution. According to

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

the authors, the study's main limitation is found to be the complexity in isolating the training effects from the practice and exercising ones. The study by Haw (2011) focuses on the differences in the psychological critical thinking as measured through the PCTE (Lawson, 1999) in two groups of students enrolled in their second and fourth year of Psychology, respectively. Second year students, during one of their programmed courses have received, together with the subject's contents, specific instructions on critical thinking: instructions were delivered by presenting them with scenarios and group discussions. The first measurement of critical thinking was made at the beginning of the semester and, as expected by the authors, shows that fourth year students obtain a higher score than first year's; the end of semester results highlight a significant improvement for second year students and no improvement for fourth year students. The authors highlight the value of specific instructions and structured exercises about critical thinking to promote psychological critical thinking. The lack of improvement in fourth year students is, according to the authors, related to the lack of homework and exercises explicitly dedicated to critical thinking. The primary limitation of the study, accordion to the authors, is related to the lack of any control group. The study by Penningroth et al. (2007) is a near-experiment involving two groups of psychology students: one group received explicit instructions on critical thinking through the Active Learning mode (Bonwell & Eison, 1991), while the other group did not receive specific instructions; both groups used the same text which included a part dedicated to critical thinking. The groups were asessed with the PCTE tool, before and at the end of the course, detecting a significant difference of critical thinking improvement in the group that received explicit instructions. A similar system was implemented in the study by Solon (2007), that is a near-experiment that involve a group of students who had to work on exercises and tasks regarding critical thinking, while the other group, that used the same text which included a chapter dedicated to critical thinking, did not receive explicit instructions about critical thinking. As for the Penningroth et al. (2007) study, Solon (2007) also detects a significant critical thinking improvement in the group that received explicit instructions, and no improvement in the other one; however, Solon (2007) has made use of Cornell Z (Ennis et al., 1985; 2004), an instrument that evaluates critical thinking in general and non specifically in the psychology field. The study by Solon (2007), moreover, assessed the specific knowledge of the course in the two groups before and after intervention, and did not find significant differences: this result, according to the authors, confirms that activities dedicated to critical thinking do not reduce the opportunity to

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

learn specific knowledge. Also, the study by Muehlenkamp et al. (2015) is a near experiment which compares two groups of psychology students, one involved in activities of promotion of critical thinking through the PBL methodology and the other without any activity explicitly directed towards critical thinking. The results highlight significant changes in favour of the first group for the different components of critical thinking identified by the authors (high level critical thinking, processes, tools), assessed through self-report tools and the application of the principles of psychological research in critical thinking, as well as through PCTE (Lawson, 1999); furthermore, as measured with the SCEQ College Student Course Engagement (Handelsman et al., 2005) tool, the study highlights a greater involvement in students who participated in the PBL training.

For both studies, the emphasized limitation is related to the absence of a control group. The work by Bensley et al. (2010) is a quasi-experiment that involved 3 groups of psychology students attending a course of psychology research methodology; for each participant the authors assessed the critical thinking ability through the CT Critical Thinking test (Bensley & Baxter, 2006), the disposition to critically think through the Need for Cognition scale (Cacioppo & Petty, 1982), the general as well as the psychology-specific academic assessments, and through scholastic assessment tests (SAT). The procedure involved the experimental group associating the research methodology course syllabus with a series of exercises explicitly dedicated to teaching critical thinking, drawing inspiration from the text by Bensley (1998), while other groups, while using the same text, were not receiving any explicit instructions. The data suggest that the number of total psychology courses attended by each student may be correlated to the initial value of critical thinking, but the improvement in critical thinking is significant only in the group receiving an explicit and direct critical thinking training. Generally speaking, the statistical screenings that were carried out appear to confirm that is, in fact, training the element that explains the change in critical thinking, and not the initial differences regarding the variables under consideration. The study by Campbell and Oswald (2018) involved psychology students engaged in home visiting support projects for children with disabilities; the students have been assessed before and after critical thinking training, through the Problem-Solving Analysis Protocol (P-SAP) (Fitch & Steinke, 2013).

Furthermore, qualitative observations were made with regards to students' participation and satisfaction in terms of educational methods. The project consisted in associating the home visiting program, that was part of the psychology course, with three types of activities: group

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

discussions, supporting and reflecting on ways to promote problem solving and critical thinking, hence guided writing tasks and analysis of scenarios. Observations made in the classroom demonstrated that students were interested in participating in the activity, as considered useful and applicable in home-visiting jobs as well. Quantitative data also show a significant change in critical thinking for all those who participated in the study, highlighting the importance of such activities associated with practical experiences.

The study by Cammeo and colleagues (2022), conducted immediately after the Covid-19 pandemic, introduces the innovative element of technology through podcasts. Specifically, the group of attending students was exposed to listening to 10 podcasts containing clinical cases of minors with various neurodevelopmental disorders (including autism, learning disorder, ADHD), while the group of non-attending students studied the same scripts in a traditional paper-based mode. The results obtained from the non-parametric statistical analysis between the pre and post-tests of both groups show an improvement in all 3 variables considered: information gathering, hypothesis generation and treatment management (Daniel et al., 2019) in favour of the group of attending students who were therefore exposed to the podcast training. Finally, the study included an analysis of the percentage of agreements and disagreements in the different categories of 3 trained trainees who independently evaluated the tests.

None of the studies measured or analyzed the satisfaction of the educational method used. Every study described medium/long-term interventions (between 12/17 weeks and 12 months) in experimental or quasi-experimental projects (with control groups and pre/post-tests).

Conclusion

Critical thinking and clinical reasoning are fundamental in most healthcare professions, however it is complex to transfer these "immaterial" abilities through traditional education tested on theoretical or technical contents (Payan-Carreira et al., 2019): the collected studies present methodologies that, although different, promote significantly the improvement of critical thinking in university students, demonstrating that, beyond the limitations of the studies, it "is possible" to transfer this complex skill. The improvement of critical thinking or one of its components, namely, as considered in this review, the skill set that leads to a reflective process that translates into interpretation, analysis and evaluation of data, in order to formulate a judgement, is promoted in all the studies analysed, highlighting how explicit teaching strategies can

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

lead to the acquisition of complex, effective and functional skills, as well as to their transferability to work contexts.

The review of the study analysed to promote critical thinking shows that different methodologies have been used (Problem Based Learning; Collaborative Approach; Active Learning; teaching practices based on text-books educational design; review of the literature, design and research methodology; writing tasks; group discussions; homework; class tests; tools as podcasts). Such methodologies focused on different contents (theoretical contents, research data interpretation, scenarios) and, finally, have been associated with different courses (developmental psychology, research methods, etc.). Despite the methodology and content dissimilarities, results have emphasized a significant increase in critical thinking or related skills competencies. Moreover, critical thinking, as defined and measured in the studies under analysis, appears to improve regardless of learning styles, motivations, previous knowledge and academic performance. Despite the methodological limitations, mainly related to the absence of control groups and lack of follow-ups, the effect of training on the critical thinking variable seems to be evident and significant in the studies that utilised self-report tools as well as in those that used other tools.

Collecting the elements in common, it is possible to draw this series of reflections: in the works analysed the improvement in critical thinking is attributed to explicit teaching methods; the studies suggest the use of a structured program to teach critical thinking, and use educational methods that include group activities and active participation of the student.

The scarcity of the studies we found for this review does not allow us to present any general reflections; therefore, we will present some considerations that may be used as background for future research. A primary element, noticeable because related to job placement competencies, concerns the fact that issues related to generalization and transferability were not tackled in any of the articles analysed; thus, a further complexity emerges with regards to assessing the efficiency of learning strategies designed to improve critical thinking and clinical reasoning.

Most of the articles only focused on the main teaching strategy used by professors or researchers, albeit such strategy is, in all of the works, a combination of different strategies which were not always described in an exhaustive way and, in almost every case, did not present an education design logic and a theoretical framework (Bensley et al., 2010; Campbell & Oswald, 2018; Haw, 2011; Karantzas et al., 2013;

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Muehlenkamp et al., 2015; Penningroth et al., 2007; Solon, 2007; Wentworth & Whitmarsh, 2017; Cammeo et al., 2022).

For instance, most studies talk about the "use of scenarios" which are not always described in detail and lack of explicit analysis of the nature of the issues or the cases to solve, as well as a guideline that allows the reader to understand the choice of using one scenario rather than another (Muehlenkamp et al., 2015; Wentworth & Whitmarsh, 2017).

Another consideration deriving from this review can be made in regards to the variety of assessment tools (tests vs. indices or surveys; standardized vs. non-standardized) used to quantify the improvements in learning the competencies studied. In accordance with Chan (2016) and Payan-Carreira (2019), every article has focused on the results from the final assessment and did not monitor changes during the learning process.

Standardized tests for critical thinking evaluation were the most used and those that showed a greater change in the pre/post test phase, while tests used to assess more general domains showed less significant margins for change.

In some articles tests were used in combination with non-standardized material (surveys or questionnaires, for instance), making the generalization of results even more complex. All the articles included in our review assessed the statistical significance (P value) as a measure of the efficacy of critical thinking learning strategies. Another complexity in the interpretation of results is related to the lack of adequate inspections and methodological strictness, given that all studies use quasi-experimental designs with practical examples, or descriptive designs without control groups, and involve only one pre-test and post-test evaluations. As suggested by previous studies (Behar-Horenstein & Niu 2011; Lapkin et al. 2010; Oliveira et al. 2016; Payan-Carreira et al., 2019), future research would have to engage in planning experimental designs which include random assignments inside the groups and, possibly, measures of change during the training. The study by Cammeo and colleagues (2022) introduces teaching methodologies that make use of technology. This strand could be useful in the planning of courses and teaching materials that can also be used at a distance.

This review is intended as a basis for designing researches aimed at investigating the impact of the different methodologies as well as a starting point to urge reflections on the possibility of including critical thinking education within the different training levels dedicated to psychologists.

Continuing on findings, this review suggests that learning strategies that actively involve students might be of preferable to traditional

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

lectures for the promotion of critical thinking and clinical reasoning. However, the presence of a limited number of studies associated to the lack of a solid theoretical background and the diversity of evaluation tools, compromise the comparison of the efficacy of the described learning activities. Therefore, using the studies' limitations as practical suggestions for future studies, we may advice the following: studies will have to be carefully designed from the methodological point of view; will have to be described in a way that is more operationalized and replicable in terms of the roles of facilitator, eventual learning tutors and students; and it will have to explain the type of materials in use, their sequence and presentation modes. Further research should concern: threats to the internal validity (for example, by at least opting for quasi-experimental projects with randomized sampling), the use of larger samples; being cautious in terms of controlling variables, such as age, gender, academic performance, pedagogical knowledge of the teachers. It is also fundamental that future studies gather information on transferability or generalization of improvements in the workplace or in the structures where the student moves after graduating.

As a starting point for professors and practitioners who work on training clinical psychologists, there is plenty of room for innovation: every scenario in use is read or told by the professor or the facilitators, and they do not implement technological solutions, such as podcasting, videos, or artificial intelligence systems that could make the situation more realistic and, perhaps, enhance learning. Furthermore, it would be useful to include qualitative surveys to explore what are the salient variables in the clinical reasoning process from the students' point of view and to extend the interventions to other professionals in order to work in a systematic way.

We believe that research should continue in this direction, because agencies interested in health care professional training (i.e. Professional Associations, Universities, Training Agencies, CME Providers and psychotherapy schools), as well as especially public and private health-care facilities that welcome postgraduate psychology trainees and/or are interested in developing professional collaborations, expect that graduates not only master the scientific and technical core knowledge of the profession, but also possess advanced thinking capabilities, that allow them to engage in clinical reasoning processes that are essential in the health and care sector (Aglen, 2016; Cavallini et al., 2019; Hildenbrand & Schultz, 2012).

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

References

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D..(2008). Instructional Interventions Affecting Critical Thinking Skills and Dispositions: A Stage 1 Meta-Analysis. *Review of Educational Research, 78*(4), 1102-34. DOI: 10.3102/0034654308326084.

Aglen, B. (2016). Pedagogical strategies to teach bachelor students evidence-based practice: A systematic review. *Nurse education today, 36*, 255-263. DOI: 10.1016/j.nedt.2015.08.025.

Ahlfeldt, S., Mehta, S., & Sellnow, T. (2005). Measurement and analysis of student engagement in university classes where varying levels of PBL methods of instruction are in use. *Higher Education Research & Development, 24*(1), 5-20. DOI: 10.1080/0729436052000318541.

Al Rumayyan, A., Ashmed, N., Al Subait, R., Al Ghamdi, G., Mahzari, M. M., Mohamed, T.A., Rotgans, J. I., Donmez, M., Mamede, S., & Schmidt, H. G. (2018). Teaching clinical reasoning through hypothetico-deduction is (slightly) better than self explanation in tutorial groups: An experimental study. *Perspective on Medical Education, 7*(2), 93-99. DOI: 10.1007/s40037-018-0409-x.

Andrews, J. J. W., Syeda, M. M. (2016). Clinical reasoning in School Psychology: From Assessment to Intervention. *Canadian Journal of School Psychology, 32*(1), 3-15. DOI: 10.1177/082957351665993.

Barrows, H.S. (1996). Problem-based learning in medicine and beyond: a brief overview. *New Directions for Teaching and Learning,* 68, 3-12. DOI: 10.1002/tl.37219966804.

Bensley, D. A. (1998). *Critical thinking in psychology: A unified skills approach.* Pacific Grove, CA: Thomson Brooks/Cole Publishing Co.

Bensley, D. A., & Baxter, C. (2006). *The critical thinking in psychology test. Unpublished manuscript.* Frostburg State University, Frostburg, MD.

Bensley, D. A., Crowe, D. S., Bernhardt, P., Buckner, C., & Allman, A. L. (2010). Teaching and assessing critical thinking skills for argument analysis in psychology. *Teaching of Psychology, 37*(2), 91-96. DOI: 10.1080/00986281003626656.

Behar-Horenstein, L. S., & Niu, L. (2011). Teaching critical thinking skills in higher education: A review of the literature. *Journal of College Teaching & Learning (TLC), 8*(2), 25-41. DOI: 10.19030/tlc.v8i2.3554.

Bonwell, C. C., & Eison, J. A. (1991). Active Learning: Creating Excitement in the Classroom. ASHE-ERIC Higher Education Report. School of Education and Human Development, George Washington University.

Burbach, B., Barnason, S., & Thompson, S. A. (2015). Using “think aloud” to capture clinical reasoning during patient simulation. *International journal of nursing education scholarship, 12*(1), 1-7. DOI: 10.1515/ijnes-2014-0044.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116-131. DOI: 10.1037/0022-3514.42.1.116.

Cammeo C., Prestera G., Massaro D., Marchetti A., & Cavallini, F. (2022). Teaching clinical reasoning to psychology students: an exploratory study on the potential of a podcast training. *Journal of Clinical and Developmental Psychology, 4*(2-3), 1-15. DOI: 10.13129/2612-4033/0110-3394.

Campbell, C. G., & Oswald, B. R. (2018). Promoting critical thinking through service learning: A home-visiting case study. *Teaching of Psychology, 45*(2), 193-199. DOI: 10.1177/0098628318762933.

Cavallini, F., Sacchetti, M. C., Vascelli, L., & Artoni, V. (2019). The academic network in the field of psychology as a strategy of innovation for nonprofit enterprises. *Journal of Clinical and Developmental Psychology, 1*(2), 39-48. DOI: 10.6092/2612-4033/0110-2102.

Chan, Z. C. (2016). A systematic review on critical thinking in medical education. *International journal of adolescent medicine and health, 30*(1). DOI: 10.1515/ijamh-2015-0117.

Choi, S., Oh, S., Lee, D. H., & Yoon, H. S. (2020). Effects of reflection and immediate feedback to improve clinical reasoning of medical students in the assessment of dermatologic conditions: a randomised controlled trial. *BMC Medical Education, 20*, 146. DOI: 10.1186/s12909-020-02063-y.

Delany, C., Kameniar, B., Lysk, J., & Vaughan, B. (2020). "Starting from a higher place": linking Habermas to teaching and learning clinical reasoning in the emergency medicine context. *Advances in Health Sciences Education, 25*(4), 809-824. DOI: 10.1007/s10459-020-09958-x.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In Dillenbourg, P. (Ed.), *Collaborative-Learning: Cognitive and Computational Approaches.* Oxford: Elsevier, (pp. 1-19).

Duncker, K. (1926). A qualitative (experimental and theoretical) study of productive thinking solving of comprehensible problems. *Pedagogical Seminary and Journal of Genetic Psychology, 33*, 642-708. DOI: 10.1080/08856559.1926.10533052.

Durning, S. J., Trowbridge, R. L., & Schuwirth, L. (2019). *Clinical Reasoning and Diagnostic Error. A Call to Merge Two Worlds to Improve Patient Care.* Academic Medicine: Volume Publish Ahead of Print Issue. DOI: 10.1097/ACM.0000000000003041.

Ennis, R. H., Millman, J., & Tomko, T. N. (1985). Cornell Critical *Thinking Tests, Level X and Level Z manual (3th Eds)*. Pacific Grove, CA: Midwest Publications.

Ennis, R. H., Millman, J., & Tomko, T. N. (2004). Cornell Critical *Thinking Tests, Level X and evel Z manual (4th Eds)*. Seaside, CA. Critical Thinking Company.

Faucher, C. (2011). Differentiating the Elements of Clinical Thinking. *Optometric Education, 36*(3), 140-145. https://journal.opted.org/articles/Volume_36_Number_3_CriticalThinking.pdf.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Fitch, P., & Steinke, P. (2013). *Tools for assessing cognitive outcomes of experiential learning. NCA HLC Annual Conference on Quality in Higher Education.* https://departments.central.edu/psychology/files/2011/07/HLC-2013-Paper-Fitch-Steinke.pdf.

Guerrero, J. (2019). Practice Rationale Care Model: The Art and Science of Clinical Reasoning, Decision Making and Judgment in the Nursing Process. *Open Journal of Nursing, 9*, 88. DOI: 10.4236/ojn.2019.92008.

Handelsman, M. M., Briggs, W. L., Sullivan, N., & Towler, A. (2005). A measure of college student course engagement. *The Journal of Educational Research, 98*, 184-192. DOI: 10.3200/JOER.98.3.184-192.

Haw, J. (2011). Improving psychological critical thinking in Australian university students. *Australian Journal of Psychology, 63*(3), 150-153. DOI: 10.1111/j.1742-9536.2011.00018.x.

Hildenbrand, K. J., & Schultz, J. A. (2012). Development of a rubric to improve critical thinking. *Athletic Training Education Journal, 7*(3), 86-94. DOI: 10.5608/070386.

Huhn, K., Gilliland, S. J., Black, L. L., Wainwright, S. F., & Christensen, N. (2019). Clinical Reasoning in Physical Therapy: A Concept Analysis. *Physical Therapy, 99*(4), 440-456. DOI: 10.1093/ptj/pzy148.

Ihm, J., Shin, Y., & Seo, D. G. (2020). Did Clinical reasoning and Knowledge Questions During Team-based Learning Enhance Dental Students' Performance in Esthetic Dentistry?. *Journal of Dental Education, 84*(4), 495-501. DOI: 10.21815/JDE.019.19.

Joice, V. C. (2013). Critical Thinking Versus Clinical Reasoning Versus Clinical Judgment: Differential Diagnosis. *Nurse Educator, 38*(1), 34-36. DOI: 10.1097/NNE.0b013e318276dfbe.

Karantzas, G. C, Avery, M. R., Macfarlane, S., Mussap, A., Tooley, G., Hazelwood, Z., & Fitness, J. (2013). Enhancing critical analysis and problem-solving skills inundergraduate psychology: An evaluation of a collaborative learning and problem-based learning approach. *Australian Journal of Psychology, 65*, 38-45. DOI: 10.1111/ajpy.12009.

Kosior, K., Wall, T., & Ferrero, S. (2019). The role of metacognition in teaching clinical reasoning: Theory to practice. *Education in The Health Professions, 2*(2), 108-114. DOI: 10.4103/EHP.EHP_14_19.

Koufidis, C., Manninen, K., Nieminen, J., Wohlin, M., & Silén, C. (2020). Grounding judgement in context: A conceptual learning model of clinical reasoning. *Medical education, 54*(11), 1019-1028. DOI: 10.1111/medu.1422.

Lapkin, S., Levett-Jones, T., Bellchambers, H., & Fernandez, R. (2010). Effectiveness of patient simulation manikins in teaching clinical reasoning skills to undergraduatenursing students: A systematic review. *Clinical simulation in nursing, 6*(6), e207-e222. DOI: 10.1016/j.ecns.2010.05.005.

Lawson, T. J. (1999). Assessing psychological critical thinking as a learning outcome for psychology majors. *Teaching of Psychology, 26*, 207-209. DOI: 10.1207/S15328023TOP260311.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial –
No Derivatives License. For terms and conditions of usage
please see: http://creativecommons.org

Li, S., Eang, Y., Xue, J., Zhao, N., & Zhu, T. (2020). The Impact of COVID-19 Epidemic Declaration on Psychological Consequences: A Study on Active Weibo Users. *International Journal Environmental Research and Public Health, 17*(6), 2032. DOI: 10.3390/ijerph17062032.

Mason, A., Locke, R., Lusznat, R. M., Coles, C., & Masding, M. G. (2020). How Do Contexts Affect Physicians' Clinical Reasoning? A Narrative Review. *MedEdPublish, 9*(1), 32. DOI: 10.15694/mep.2020.000032.1.

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research, 14*, 579. DOI: 10.1186/s12913-014-0579-0.

Ministero della Salute (2018) Rapporto salute mentale. Analisi dei dati del Sistema informativo per la salute mentale (SISM). Available online: http://www.salute.gov.it/portale/documentazione/p6_2_2_1.jsplingua=italiano&id=2932.

Muehlenkamp, J. J., Weiss, N., & Hansen, M. (2015). Problem-based learning for introductory psychology: Preliminary supporting evidence. *Scholarship of Teaching and Learning in Psychology, 1*(2), 125. DOI: 10.1037/stl0000027.

Mutter, M. K., Martindale, J. R., Shah, N., Gusic, M. E., & Wolf, S. J. (2020). Case-Based Teaching: Does the Addition of High-Fidelity Simulation Make a Difference in Medical Students' Clinical Reasoning Skills?. *Medical Science Educator, 30*, 307-313. DOI: 10.1007/s40670-019-00904-0.

Norman, G. R. (2000). The epistemology of clinical reasoning: perspectives from philosophy, psychology, and neuroscience. *Academic Medicine, 75*(10), S127-S133. DOI: 10.1097/00001888-200010001-00041.

O'Hare, L., & McGuinness, C. (2004). Skills and attributes developed by psychology undergraduates: Ratings by undergraduates, post graduates, academic psychologists and professional practitioners. *Psychology Learning and Teaching, 4*, 35-42. DOI: 10.2304/plat.2004.4.1.35.

Oliveira, L. B. D., Díaz, L. J. R., Carbogim, F. D. C., Rodrigues, A. R. B., & Püschel, V. A. D. A. (2016). Effectiveness of teaching strategies on the development of critical thinking in undergraduate nursing students: a meta-analysis. *Revista da Escola de Enfermagem da USP, 50*(2), 355-364. DOI: 10.1590/S0080-623420160000200023.

Payan-Carreira, R., Cruz, G., Papathanasiou, I. V., Fradelos, E., & Jiang, L. (2019). The effectiveness of critical thinking instructional strategies in health professions education: a systematic review. *Studies in Higher Education, 44*(5), 829-843. DOI: 10.1080/03075079.2019.1586330.

Pelaccia, T., Plotnick, L. H., Audétat, M. C., Nendaz, M., Lubarsky, S., Torabi, N., Thomas, A., Young, M., & Dory, V. (2020). A Scoping Review of Physicians' Clinical Reasoning in Emergency Departments. *Annals of Emergency Medicine, 75*(2), 206-217. DOI: 10.1016/j.annemergmed.2019.06.023.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Penningroth, S. L., Despain, L. H., & Gray, M. J. (2007). A course designed to improve psychological critical thinking. *Teaching of Psychology, 34*(3), 153-157. DOI: 10.1080/00986280701498509.

Pierce, C., Corral, J., Aagaard, E., Harnke, B., Irby, D. M., & Stickrath, C. (2020). A BEME realist synthesis review of the effectiveness of teaching strategies used in the clinical setting on the development of clinical skills among health professionals: BEME Guide No.61 [published online ahead of print, 2020 Jan 21]. *Medical Teacher*, 1-12. DOI: 10.1080/0142159X.2019.1708294 .

Pramila-Savukoski, S., Juntunen, J., Tuomikoski, A. M., Kaariainen, M., Tomietto, M., Kaucic, B. M., Filej, B., Riklikiene, O., Vizcaya-Moreno, M. F., Perez-Canaveras, R. M., De Raeve, P., & Mikkonen, K. (2019). Mentors' self-assessed competence in mentoring nursing students in clinical practice: A systematic review of quantitative studies. *Journal of Clinical Nursing, 29*(5-6), 684-705. DOI: 10.1111/jocn.15127.

Richmond, A., Cooper, N., Gay, S., Atiomo, W., & Patel, R. (2020). The student is key: A realist review of educational interventions to develop analytical and non-analytical clinical reasoning ability. *Medical education, 54*(8), 709-719. DOI: 10.1111/medu.14137.

Robertson, D. M., (2012). *Critical thinking and clinical reasoning in new graduate occupational therapists: a phenomenological study*. Available from OpenAIR@RGU. [online]. Available from: http://openair.rgu.ac.uk

Roschelle, J., & Teasley, S. (1995). The construction of shared knowledge in collaborative problem solving. In C. E. O'Malley (Eds.), *Computer supported collaborative learning* (pp. 69-97). Heidelberg: Springer-Verlag. DOI: 10.1007/978-3-642-85098-1_5.

Frank, J. R., Snell, L., & Sherbino J. CanMEDS (2015). *Physician Competency.* Framework. Royal College of Physicians and Surgeons of Canada (2015). Available online:
http://www.royalcollege.ca/rcsite/canmeds/canmeds-framework-e.

Royce, C. S., Hayes, M. M., & Schwartzstein, R. M. (2019). Teaching Critical Thinking: A Case for Instruction in Cognitive Biases to Reduce Diagnostic Errors and Improve Patient Safety. *Academic Medicine, 94*(2), 187-194. DOI: 10.1097/ACM.0000000000002518.

Rutter, P. M., & Harrison, T. (2020). Differential diagnosis in pharmacy practice: Time to adopt clinical reasoning and decision making. *Research in Social and Administrative Pharmacy, 16*(10), 1483-1486. DOI: 10.1016/j.sapharm.2020.02.020.

Shin, H. S. (2019). Reasoning processes in clinical reasoning: from the perspective of cognitive psychology. *Korean Journal of Medical Education, 31*(4), 299-308. DOI: 10.3946/kjme.2019.140.

Solon, T. (2007). Generic Critical Thinking Infusion and Course Content Learning in Introductory Psychology. *Journal of Instructional Psychology, 34*(2), 95-109.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

Sternberg, R. J., Halpern, D. F. (2020). *Critical Thinking in Psychology, 2th (Eds). Promoting Critical Thinking by Teaching, or Taking, Psychology Courses*, 5, pp. 103-124. Cambridge University Press.

Cumming, A., Cumming, A., & Ross, M. (2007). The Tuning Project for Medicine-learning outcomes for undergraduate medical education in Europe. *Medical teacher, 29*(7), 636-641. Consulted on 12 January 2020. DOI: 10.1080/01421590701721721.

Torre, D., Durning, S. J., Rencic, J., Lang, V., Holmboe, E., & Daniel, M. (2020). Widening the lens on teaching and assessing clinical reasoning: from "in the head" to "out in the world". *Official Journal of the Society to Improve Diagnosis in Medicine (SIDM), 7*(3), 181-190. DOI: 10.1515/dx-2019-0098.

Vidyarthi, A., Lek, N., Chan, K., & Kamei, R. (2016). Experiences with a clinical reasoning and evidence-based medicine course. *The clinical teacher, 13*(1), 52-57. DOI: 10.1111/tct.12362.

Wentworth, D. K., & Whitmarsh, L. (2017). Thinking Like a Psychologist Introductory Psychology Writing Assignments: Encouraging Critical Thinking and Resisting Plagiarism. *Teaching of Psychology, 44*(4), 335-341. DOI: 10.1177/0098628317727909.

Young, M., Thomas, A., Gordon, D., Gruppen, L., Lubarsky, S., Rencic, J., Ballard, T., Holmboe, E., Da Silvia, A., Ratcliffe, T., Schuwirth, L., & Durning, S. J. (2019). The terminology of clinical reasoning in health professions education: Implications and considerations. *Journal of Medical Teacher, 41*(11), 1277-1284. DOI: 10.1080/0142159X.2019.1635686.

Young, M. E., Thomas, A., Lubarsky, S., Gordon, D., Gruppen, L. D., Rencic, J., Ballard, T., Holmboe, E., Da Silva, A., Ratcliffe, T., Schuwirth, L., Dory, V., & Durning, S. J. (2020). Mapping clinical reasoning literature across the health professions: a scoping review. *BMC Medical Education, 20*, 107. DOI: 10.1186/s12909-020-02012-9.

Copyright © FrancoAngeli
This work is released under Creative Commons Attribution - Non-Commercial – No Derivatives License. For terms and conditions of usage please see: http://creativecommons.org

RICERCHE DI PSICOLOGIA

Ricerche di Psicologia è una rivista scientifica che pubblica articoli di tipo teorico, storico, empirico e di rassegna e discussione critica che affrontano tutti i temi della psicologia. È di interesse della rivista ogni argomento della psicologia, considerato sotto l'aspetto dei più recenti avanzamenti della ricerca: psicologia generale, cognitiva e sperimentale; storia della psicologia; psicologia fisiologica, neuropsicologia e psicobiologia; psicologia animale e comparata; psicometria e metodologia della ricerca psicologica; psicologia dello sviluppo e dell'educazione; psicologia sociale; psicologia del lavoro e delle organizzazioni; psicologia dinamica; psicologia clinica. La rivista è sensibile a un approccio che cerca di intercettare le questioni dell'attualità e i problemi che emergono dal mondo della cultura e dalla società, favorendo il dialogo interdisciplinare. Ricerche di Psicologia può ospitare forum centrati su specifiche tematiche o che raccolgono contributi scientifici associati a eventi e convegni e la presentazione organica dei risultati di progetti di rilevante interesse. La rivista pubblica articoli in italiano e inglese, incoraggiando i secondi per favorire la diffusione internazionale.

Edizione fuori commercio

ISSNe 1972-5620