

CODEBOOK

The following table refers to data presented in Dataset_Complete.CSV

This data has been collected between March and June 2018 and 2019 in the province of Milan (Italy).

Name of the variable	Content and meaning of the variable
ID	Identification number of each child in the sample (1-416)
Raven_Percentile	Percentile score of each participant, corrected by age (0-99)
Age_Years	Age of participants expressed in years. 3 = 3 years old child 4 = 4 years old child 5 = 5 years old child 6 = 6 years old child
Sex	Sex of each participant. 1 = male 2 = female
School	School attended by each participant (1-9)
Evaluator	Evaluator that administered the battery (A-G)
Class	Class attended by each participant (A-Z)
Age_Months	Age of participants expressed in months (36-75)
DCCS	Score of the participant at the Dimensional Change Card Sorting task (from FE-PS battery; Usai et al., 2013). Scores obtained represent the total number of correct classifications performed by the child (range: 0-24).
Statue	Score of the participant at the Statue (from Nepsy II; Korkman et al., 2011). Scores obtained represent total errors made by the child in performing the task (range: 0-45).
DigitSpan_Forward	Score of the participant at the Digit Span in the forward condition (from the WISC IV; Wechsler, 2012). Scores obtained represent the maximum list length (0 – 16) at which the sequences are correctly recalled.
DigitSpan_Backward	Score of the participant at the Digit Span in the backward condition (from the WISC IV; Wechsler, 2012). Scores obtained represent the maximum list length (0 – 16) at which the sequences are correctly recalled.
Item1-Item 55	Items 1-55 refer to the response of participants' teacher at the QUVA-p. QUVA-p is a teacher-completed screening questionnaire that measures dysregulation problems and socio-relational and adaptive issues at preschool. The questionnaire comprises 55 items, all scored on a 3-point Likert scale (0 = not at all/never, 1 = sometimes, 2 = often/always). Responses at positive items (i.e.,

	7,12,13,22,23,25,28,30,31,36,39,46,47,48,49,50,51,52,53,54,55) are reversed.
Mother_study	Index of socio-cultural status referred to mother's education level. 1 = elementary school 2 = middle school 3 = high school 4 = university
Mother_Work	Index of socio-economic status referred to mother's profession. 1 = unemployed 2 = workman 3 = technical worker 4 = management
Father_Study	Index of socio-cultural status referred to father's education level. 1 = elementary school 2 = middle school 3 = high school 4 = university
Father_Work	Index of socio-economic status referred to father's profession. 1 = unemployed 2 = workman 3 = technical worker 4 = management

Random Dataset_RQ5 is composed of data of 92 participants randomly selected from children with and without critical scores at QUVA-p in Dataset_Complete.CSV
This dataset differs from the Dataset_Complete.CSV for the presence of four new variables.

Name of new variables	Content and meaning of new variables
Group_of_risk	This variable represents each participant's risk of having or not having dysregulation issues, based on their score at the QUVA-p. 0 = not at-risk child 1 = at-risk child In order to classify each child by dysregulation risk, a total composite of dysregulation was computed (sum of the score at cognitive, behavioural, and emotional dysregulation subscales) and converted in the rank percentile corrected by age and sex. Norms for the conversion are available in Conti, Scionti and Marzocchi (2020). Children with scores above the 85 th percentile have been classified as at risk.
Emotional_Dys_Factor	The factor has been derived through Confirmatory Factor Analysis. It is indexed by items addressing emotional dysregulation issues (item 4,5,6,8,9,10,37,44,45). Factor scores are the sum of the responses at these items multiplied by factor weights.

Behavioral_Dys_Factor	The factor has been derived through Confirmatory Factor Analysis. It is indexed by items addressing behavioural dysregulation issues (item 1,2,3,7,15,26,27,29,30,31,32,33,35,40). Factor scores are the sum of the response at these items multiplied by factor weights.
SocRelat_Difficulties_Factor	The factor has been derived through Confirmatory Factor Analysis. It is indexed by items addressing socio-relational and adaptive issues (item 21,22,23,24,28,34,36,38,47,48,49,50,51,52,53,54,55). Factor scores are the sum of the response at these items multiplied by factor weights.
CognitiveDys_Factor	The factor has been derived through Confirmatory Factor Analysis. It is indexed by items addressing cognitive dysregulation issues (item 11,12,13,14,16,17,18,19,20,25,39,41,42,43,46). Factor scores are the sum of the response at these items multiplied by factor weights.