Harnessing Linguistic Analysis for ADHD Diagnosis Support: A Stylometric Approach to Self-Defining Memories

Florian Cafiero[1], Juan Barrios[2], Simon Gabay[2], Martin Debbané[2]
[1] Université Paris Sciences et Lettres
florian.cafiero@chartes.psl.eu
[2] Université de Genève
{juan.barrios,simon.gabay,martin.debbane}@unige.ch

Abstract

This study explores the potential of stylometric analysis in identifying Self-Defining Memories (SDMs) authored by individuals with Attention-Deficit/Hyperactivity Disorder (ADHD) versus a control group. A sample of 198 SDMs were written by 66 adolescents and were then analysed using Support Vector Classifiers (SVC). The analysis included a variety of linguistic features such as character 3-grams, function words, sentence length, or lexical richness among others. It also included metadata about the participants (gender, age) and their SDMs (self-reported sentiment after recalling their memories). The results reveal a promising ability of linguistic analysis to accurately classify SDMs, with perfect prediction (F1=1.0) in the contextually simpler setup of text-by-text prediction, and satisfactory levels of precision (F1 = 0.77) when predicting individual by individual. Such results highlight the significant role that linguistic characteristics play in reflecting the distinctive cognitive patterns associated with ADHD. While not a substitute for professional diagnosis, textual analysis offers a supportive avenue for early detection and a deeper understanding of ADHD.

Keywords: ADHD, psycholinguistics, NLP, stylometry

1. Introduction

The use of Natural Language Processing (NLP) methods in psychology is both useful and complex. Useful because prediction tools have been proven for years to assist clinicians in their work. However, it is also complex because it is difficult to bring together enough people with a specific disorder to obtain a satisfactory corpus for a NLP experiment. This is particularly the case of Attention-Deficit/Hyperactivity Disorder (ADHD) in adolescents (cf. Barrios et al. 2023), a disorder with high prevalence in the population (±5.6% in teenagers aged 12 to 18 years, cf. Salari et al. 2023) producing a high level of impairment in daily life. In this paper, we propose to examine the question of ADHD in adolescents and, on the basis of a corpus recently collected in Geneva, to improve the diagnosis in young patients using machine learning techniques.

1.1. The impact of ADHD

ADHD is a prevalent neurodevelopmental disorder characterised by differences in large-scale neural connectivity (Rafi et al., 2023) and symptoms of inattention, hyperactivity, and impulsivity. Children with ADHD often struggle with impaired academic performance (Español-Martín et al., 2023), emotional regulation (Rathje et al., 2023), and different mentalisation abilities (Poznyak et al., 2023). As a result, these impairements often lead to disruptive behaviours, difficulties to maintain relationships, and challenges in daily functioning (Barkley, 2015). If not addressed in time, these impairments can extend from adolescence to adulthood, contributing to academic underachievement, substance abuse, and mental health problems such as depression and anxiety (Faraone et al., 2024). It is therefore crucial to detect ADHD to tackle these psychosocial obstacles and difficulties as soon as possible, by providing support within educational settings and nurturing the growth of social skills, in order to promote positive development and improve the well-being of people with ADHD. (Barkley, 2015).

1.2. The detection of ADHD

The early diagnosis of ADHD is fundamental (Wolraich et al., 2019), but it is both complex and time consuming, especially because of the comorbidities that can co-occur with ADHD or that can mimic similar symptoms to ADHD (Barkley, 2014). Furthermore, it mostly relies on subjective evaluations of observed behaviours, which can produce biases during psychological assessment and for differential diagnoses (Miyasaka et al., 2018). The use of computerised tests to incorporate objective data in the assessment process has already been proposed to address the aforementioned issues (Gualtieri and Johnson, 2005), but alternative techniques always have to be tested.

The advent of NLP and stylometric methods presents new avenues for computerised assessment technology, enabling the generation of rich objective data to improve the diagnosis. By quan-

Proceedings of the RaPID-5 @LREC-COLING 2024, pages 87–94
21 May, 2024.

titatively analysing the linguistic and stylistic features of texts written by diagnosed persons, researchers can uncover linguistic fingerprints that traditional methods may miss (Cafiero and Camps, 2022). Previous research has demonstrated notable linguistic differences between individuals with ADHD and control groups (Yoder, 2006; Kim and Lee, 2009; Kim et al., 2015). However, the idea of using these differences to assist psychologists and psychiatrists in diagnosing ADHD using traditional stylometric methods (Barrios et al., 2023) is new.

1.3. Understanding Psychopathological Processes through Narrative

Integrating narrative approaches to psychopathology (Lind et al., 2022) is a field in constant growth (Waters and Fivush, 2015; Adler et al., 2016; Vanden Poel and Hermans, 2019; Reed et al., 2020). It specifically studies how people make sense of their life's experiences and how the resulting script of their life changes according to changing conditions, the evolution of their main goals, etc. The profound paradox of this process of meaning-making lies in the fact that although individuals change in their ways of being and living, they remain recognisable as the same person, which remains in a certain sense unchanged – a paradox very similar to that of authorial attribution, according to which a person's literary style remains stable despite stylistic changes over a lifetime.

Such a narrative approach implies the existence of written or oral linguistic material, the usage of which (i.e., the choice of pronouns or function words, etc.) provides significant understanding of someone's psychological state (Tausczik and Pennebaker, 2009; Pennebaker et al., 2003; Pennebaker and King, 1999). This insight is pivotal, as these identifiable linguistic patterns can serve as a tool to help diagnose mental health, highlighting the importance of language on understanding and assessing mental well-being. Among the different types of narratives that can be used, Autobiographical Memories (AM) are a particularly important resource.

Indeed, AM encompass memories of personal experiences and events, and therefore serve as the foundation for constructing our life narratives and ultimately the main script of our life story. Within the framework of AM, certain memories known as Self-Defining Memories (SDMs) are of particular significance. SDMs refer to events that are highly relevant to identity processes (Singer et al., 2007; Blagov and Singer, 2004), characterised by their vividness, emotional intensity, frequent recall and focus on the individual persistent concerns or unresolved conflicts (Singer et al., 2012). As such, they are the building blocks of an individual's life story and are essential to form a coherent and continuous sense of self (Conway and Pleydell-Pearce, 2000). In fact, while recalling and reflecting on SDMs, individuals construct narratives that highlight meaningful life events, significant relationships, and / or central values (Singer and Blagov, 2004). It is this repeated retrieval and reinterpretation over time that reinforces certain aspects of identity, potentially reshaping others (McAdams, 2013; Bluck and Alea, 2002), and can influence an individual's self-concept, its worldview, or emotional well-being (Berntsen and Rubin, 2006).

During the transition from adolescence to early adulthood, the emergence of SDMs marks a pivotal phase in psychological development. In this period of life, people actively engage in identity exploration and self-reflection, constructing narratives that shape their sense of self and their experiences (McAdams, 2013). This phase is of particular significance for understanding psychopathology, as disturbances in identity formation and autobiographical memory can contribute to various mental health issues (Branje et al., 2021). Therefore, a comprehensive analysis of the content, structure, and patterns of narratives during this period provides insight into identity development and can be used as a window into psychological processes (Conway and Pleydell-Pearce, 2000; Berntsen and Rubin, 2006; Singer et al., 2007; McAdams, 2013) as well as data for the automatic detection of ADHD.

2. Method

2.1. Participants

66 adolescents (15.55 ± 1.78 years; 25 men and 31 women) were included in the experiment (cf. tab. 1). Adolescents with ADHD were recruited through advertisements in local parents' associations for children with ADHD and through collaborations established with local child psychiatrists. Participants in the control group were recruited by undergraduate students attending the Faculty of Psychology and Sciences of Education at Geneva University, Switzerland. The inclusion criteria for all the participants were age (12-17 years), fluency in French, and, for the ADHD group, meeting current diagnostic criteria for ADHD (DSM-V, American Psychiatric Association, 2013). Non-fluent francophone speakers and individuals currently under psychiatric treatment were excluded from the study.

The diagnostic criteria for ADHD were investigated by detailed anamnestic interviews and confirmed using the "ADHD Child Evaluation" (ACE) (Young, 2015). All diagnostic assessments were conducted by experienced clinical psychologists specialised in ADHD.

The final ADHD sample meeting inclusion criteria

	ADHD	Control
Men	13 (52%)	22 (56.66%)
Women	12 (48%)	19 (46.34%)
12-15 yo	12 (48%)	9 (21.95%)
16-17 yo	13 (52%)	32 (78.05%)
Total	25 (100%)	41 (100%)

Table 1: Description of participants

consisted of 25 participants, 16 were diagnosed with the inattentive modality of ADHD, 1 with the hyperactive modality, while 8 exhibited the mixed modality of ADHD.

2.2. Data

The SDMs were collected using the Self-Defining Memories task (Singer and Blagov, 2001; Thorne and McLean, 2001). Following its procedure, participants were asked to evoke personal memories of events (the SDMs) meeting six criteria: they (1) occurred at least one year ago and were (2) important and generally vividly represented; (3) meaningful and useful to help themselves or a significant other understand who they are; (4) were related to an important and enduring theme and linked to other events on the same topic; (5) were either positive or negative and generate strong feelings; and finally, (6) were recalled many times.

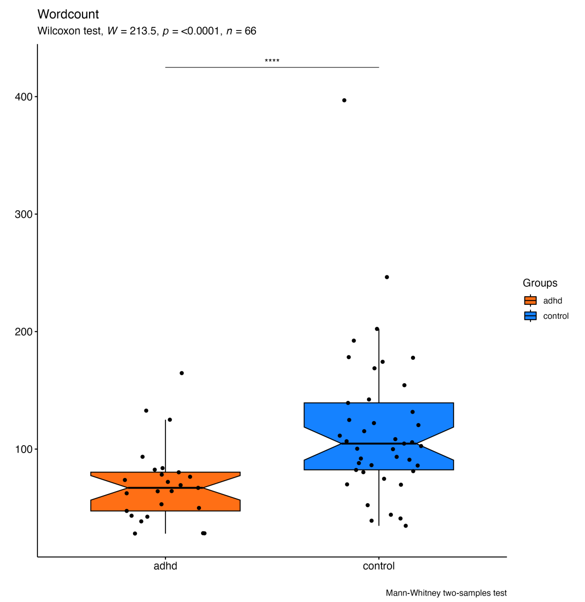

Figure 1: Number of tokens per SDM in ADHD group vs control group.

Participants were then told to imagine a situation where they met someone they liked very much and with whom they agreed, during a walk, to talk about who they really are, their “Real Me”, sharing several personal past events that powerfully convey how they have become the person they currently are. Participants were given three sheets of paper on which they had to write down, on each sheet, one SDM with a one-sentence summary. The SDMs were then transcribed by researchers, and the spelling corrected on the fly[1]. It is important to note that the SDMs produced by the two groups are quite different, particularly in terms of length (cf. fig. 1).

	ADHD	**Control**
SDMs	75	123
Positive Affect		
0	20 (26.67%)	27 (21.95%)
1	1 (1.33%)	5 (4.07%)
2	0 (0%)	7 (5.69%)
3	4 (5.33%)	7 (5.69%)
4	4 (5.33%)	8 (6.5%)
5	11 (14.67%)	18 (14.63%)
6	35 (46.67%)	51 (41.46%)
Negative Affect		
0	46 (61.33%)	50 (40.65%)
1	8 (10.67%)	11 (8.94%)
2	3 (4%)	13 (10.57%)
3	1 (1.33%)	3 (2.44%)
4	5 (6.67%)	12 (9.76%)
5	6 (8%)	9 (7.32%)
6	6 (8%)	25 (20.33%)

Table 2: Description of SDMs per group

Thereafter, participants were asked to rate their feelings after recalling each SDM on a 7-point rating scale from 0 (: not at all) to 6 (: extremely). The score distribution follows a U-shape (cf. tab. 2), which implies a tendency to score affects at the extremes, and the values do not correlate with the two groups according to a χ^2 analysis.

2.3. Textual profiling

2.3.1. Feature extraction

To predict if a text has been written by an adolescent diagnosed with ADHD or not, we train classifiers on a variety of linguistic features (character 3-grams, words, words bigrams, function words, type token ratio, text length, average sentence length) that have been proven reliable features by previous literature (Barrios et al., 2023), as well as with outputs from our experiment (text length, type/token ratio...) and information about the participant (age, gender...).

This consolidated feature matrix serves as the input to a machine learning pipeline, at the heart

[1] This procedure was implemented before starting the computational experiments.

of which lies a Support Vector Classifier (SVC), with a grid search for optimal kernel (linear, polynomial, sigmoid, RBF) and hyper parameters (cost C and γ when relevant). This choice of classifier aligns with the high-dimensional nature of the feature space and is well-regarded for text classification tasks in recent articles and surveys (HaCohen-Kerner, 2022; Bevendorff et al., 2023; Fauzi et al., 2023), and particularly fit in the case of shorter texts (Cafiero and Camps, 2021, 2023; Vogel and Meghana, 2021; Suresh Kumar et al., 2024).

2.3.2. Model Evaluation

Model robustness and generalisability are assessed through Leave-One-Out Cross-Validation (LOOCV), an evaluation method that iterates over the dataset, using each document once as a test instance while training on the remainder. Such an approach ensures that every document contributes to the validation process, which is critical in scenarios with limited data such as ours, that hardly allow other methods such as K-fold cross validation.

To avoid overfitting, we run a grid search on the C parameter, and check if the models hold when it is set to low values. The model is thus encouraged to find a hyperplane with a larger margin, which can lead to better generalisation on unseen data, at the cost of possibly underfitting the training data.

3. Results

3.1. Classification of individual texts of Self Defining Memories

Classifying texts is paradoxically the easiest task in our context, as it triples our data points (each person has written 3 SDMs) in a relatively small database. The quality of the results holds even for very low values of the C parameter (0.01).

	Precision	Recall	F1	Support
ADHD	1.00	1.00	1.00	75
Control	1.00	1.00	1.00	123
Accuracy			1.00	198
Macro avg	1.00	1.00	1.00	198

Table 3: SVM classification of individual SDMs: character 3-grams

3.2. Classification of individuals

In this experiment, we concatenate all three SDMs written by each participant, and try to predict if the person belongs to the ADHD group or the control group. The task is in our case counter-intuitively more complex because of the objectively important, but statistically speaking small, number of participants: the quantity of text remains the same, but the number of data points diminishes. Classifying individuals is redundant in our case, because the results are already more than satisfactory at the text level (1 individual=3 texts). But as it artificially makes the task harder, it helps us evaluate more complex models that could prove to be helpful facing unseen data.

We test three settings:

1. a purely lexical and syntactic analysis of the texts;
2. a setup purely relying on self reported affects and information;
3. a mix of the most relevant items.

For each of these settings, we test the various combinations of point of measures at our disposal.

3.2.1. Setup 1: linguistic classifier

The best setup we get according to our objectives does not rely only on character 3-grams, but concatenates character 3-grams, words, lexical richness and average sentence length as classifying features.

	Prec.	Rec.	F1	Supp.
ADHD	0.78	0.56	0.65	25
Control	0.77	0.90	0.83	41
Accuracy			0.77	66
Macro avg	0.77	0.73	0.74	66
Weighted avg	0.76	0.76	0.75	66

Table 4: SVM classification of individuals: best linguistic classifier for accuracy

It yields a satisfactory accuracy but unfortunately fails to provide a good recall for the ADHD group, which means that some texts written by adolescents with ADHD have minimally significant linguistic markers. These results could be linked to the different forms of ADHD (cf. § 2.1).

3.2.2. Setup 2: background and self-report affect

Classifying only on reported affect, be it the positive or negatives values given for each text, or an aggregated global value, are insufficient to give an accurate prediction in any combination possible. The classifier always ends up predicting one class only, even when implementing imbalance correction strategies. This indicates that the data is not sufficient in itself to predict the categories. We thus do not give a detailed report on the best models.

3.2.3. Setup 3: mixed classifier

Mixed classifiers, i.e. classifiers relying on any combination of linguistic and background information, never outperform classifiers based on linguistic features only, and provide at the very maximum the exact same performance, in terms of precision, F1 and recall, as purely linguistic classifiers alone. We thus do not give a detailed report on the best models.

4. Discussion

Regarding the satisfactory accuracy but the limited recall of linguistic classifiers for the ADHD group at the person's level, it could reflect the intricate nature of the disorder and its numerous comorbidities, and/or underline the existence of coping strategies as well as the quality of the support systems. In fact, although clinical studies have found that rates of language impairment in children with ADHD often exceed 50% (Mueller and Tomblin, 2012) and that these children present greater difficulties in expressive writing, spelling, and writing speed (Re et al., 2007), it is worth noting that some individuals with ADHD may also possess high IQ (High Intellectual Potential, cf. Tordjman et al. 2007; Rommelse et al. 2016) and excel academically and socially, which could introduce some heterogeneity in the linguistic markers of ADHD. Moreover, many people who are not diagnosed with ADHD may suffer from a variety of its symptoms to a subclinical degree, or may have ADHD and have simply not be diagnosed despite our efforts. This introduces a second source of fuzziness, this time in the control group. However, despite these inherent complexities linked to a psychological disorder, a signal is detected and warrants further investigation from a psycholinguistic point of view.

At a more general level, the implications of our findings are twofold, offering potential benefits in both clinical and linguistic domains:

1. **Enhancing Early Identification:** The ability to infer ADHD-related characteristics from textual analysis could serve as a supplementary tool for early identification of potential ADHD cases. This is particularly relevant in contexts where there is a pronounced increase in the demand for diagnostic evaluations, potentially alleviating some of the pressure on clinical services.

2. **Contributing to Psycholinguistic Insights:** By examining the nuances of language use among individuals with ADHD, our study contributes to a deeper understanding of how ADHD influences linguistic expression. This exploration not only enriches our knowledge of psycholinguistics but also opens avenues for further research into the intersection of language and psychological disorders.

Despite our promising results, it is important to state that automated text analysis for the identification of ADHD should not be viewed as a replacement for professional diagnosis. It should be envisaged as a supportive tool, that can contribute to the early detection and the understanding of ADHD through linguistic patterns.

5. Further work

A psycholinguistic analysis of the features is essential, in order to understand the linguistic particularities of ADHD. This involves not only identifying markers, which SVCs make it easier to do than LLMs, but also understanding the use of these markers. This type of analysis, however, is more likely to be done at the group level than at the text or individual level.

As our corpus is small, it is also important to obtain new data. In order to accelerate the acquisition of these, it could be useful to change method, and abandon manual writing for oral recitation, automatically transcribed with speech to text technologies. A study of the impact of such a change of medium would be interesting to carry out, in terms of quantity of data on the one hand, but also in terms of results on the other.

6. Acknowledgements

[Author 2] was funded by the Chilean National Agency for Research and Development (ANID Chile) through the PhD Abroad Scholarship, 2018 award. [Author 4] was funded by the Swiss National Science Foundation (Grant number 100014 179033), as well as the Marina Picasso Prize of the AEMD Foundation 2018.

7. Data availability

Data used in this study can be made available by the corresponding author on written request, as the authors have no legal or ethical restrictions to share the collected data in anonymised format. The code is available at the following address: `10.5281/zenodo.10953603`

8. Statement of Ethics

The clinical study protocol was reviewed and approved in 2019 by the Swiss Ethics Committee whereas the control data collection was reviewed and approved in 2015 by the University of Geneva

Ethics Committee (Faculty of Psychology and Education Sciences). Written informed consent was obtained from all participants (and their parents for underage subjects).

9. Conflict of Interest Statement

The authors have no conflict of interest to declare.

10. Bibliographical References

Jonathan M Adler, Jennifer Lodi-Smith, Frederick L Philippe, and Iliane Houle. 2016. The incremental validity of narrative identity in predicting well-being: A review of the field and recommendations for the future. *Pers. Soc. Psychol. Rev.*, 20(2):142–175.

American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders: DSM-5*, 5th ed. edition. American Psychiatric Association, Washington, DC.

Russell A Barkley, editor. 2014. *Attention-deficit hyperactivity disorder*, 4th ed. edition. Guilford Publications, New York, NY.

Russell A. Barkley. 2015. *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment*, 3th ed. edition. Guilford Press, New York - London.

Juan Barrios, Simon Gabay, Florian Cafiero, and Martin Debbané. 2023. Detecting psychological disorders with stylometry. In *Computational Humanities Research*, Paris, France. CEUR.

Dorthe Berntsen and David C. Rubin. 2006. Emotion and vantage point in autobiographical. *Cognition & Emotion*, 20(8):1193–1215.

Janek Bevendorff, Ian Borrego-Obrador, Mara Chinea-Ríos, Marc Franco-Salvador, Maik Fröbe, Annina Heini, Krzysztof Kredens, Maximilian Mayerl, Piotr Pęzik, Martin Potthast, et al. 2023. Overview of pan 2023: Authorship verification, multi-author writing style analysis, profiling cryptocurrency influencers, and trigger detection: Condensed lab overview. In *International Conference of the Cross-Language Evaluation Forum for European Languages*, pages 459–481. Springer.

Pavel Blagov and Jefferson Singer. 2004. Four dimensions of self-defining memories (specificity, meaning, content, and affect) and their relationships to self-restraint, distress, and repressive defensiveness. *Journal of Personality*, 72(3):481–511.

Susan Bluck and Nicole Alea. 2002. Exploring the functions of autobiographical memory: Why do I remember the autumn? In *Critical advances in reminiscence work: From theory to application.*, pages 61–75. Springer Publishing Company.

Susan Branje, Elisabeth L de Moor, Jenna Spitzer, and Andrik I Becht. 2021. Dynamics of identity development in adolescence: A decade in review. *J. Res. Adolesc.*, 31(4):908–927.

Florian Cafiero and Jean-Baptiste Camps. 2021. 'Psyché'as a Rosetta Stone? assessing collaborative authorship in the French 17th century theatre. *Proceedings http://ceur-ws. org ISSN*, 1613:0073.

Florian Cafiero and Jean-Baptiste Camps. 2022. *Affaires de style: du cas Molière à l'affaire Grégory: la stylométrie mène l'enquête*. Le Robert, Paris.

Florian Cafiero and Jean-Baptiste Camps. 2023. Who could be behind QAnon? authorship attribution with supervised machine-learning. *Digital Scholarship in the Humanities*, 38(4):1418–1430.

M. A. Conway and C. W. Pleydell-Pearce. 2000. The construction of autobiographical memories in the self-memory system. *Psychological Review*, 107(2):261–288.

Gemma Español-Martín, Mireia Pagerols, Raquel Prat, Cristina Rivas, Josep Antoni Ramos-Quiroga, Miquel Casas, and Rosa Bosch. 2023. The impact of attention-deficit/hyperactivity disorder and specific learning disorders on academic performance in spanish children from a low-middle- and a high-income population. *Front. Psychiatry*, 14:1136994.

Stephen V Faraone, Mark A Bellgrove, Isabell Brikell, Samuele Cortese, Catharina A Hartman, Chris Hollis, Jeffrey H Newcorn, Alexandra Philipsen, Guilherme V Polanczyk, Katya Rubia, Margaret H Sibley, and Jan K Buitelaar. 2024. Attention-deficit/hyperactivity disorder. *Nat. Rev. Dis. Primers*, 10(1).

Stephen V Faraone and Henrik Larsson. 2019. The genetics of attention deficit hyperactivity disorder. *Molecular psychiatry*, 20(1):17–28.

Muhammad Ali Fauzi, Stephen Wolthusen, Bian Yang, Patrick Bours, and Prosper Yeng. 2023. Identifying sexual predators in chats using SVM and feature ensemble. In *2023 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC)*, pages 1–6. IEEE.

C Thomas Gualtieri and Lynda G Johnson. 2005. ADHD: Is objective diagnosis possible? *Psychiatry (Edgmont)*, 2(11):44–53.

Tilmann Habermas. 2011. Autobiographical reasoning: arguing and narrating from a biographical perspective. *New Directions for Child and Adolescent Development*, 2011(131):1–17.

Yaakov HaCohen-Kerner. 2022. Survey on profiling age and gender of text authors. *Expert Systems with Applications*, 199:117140.

Kyungil Kim and Chang Hwan Lee. 2009. Distinctive linguistic styles in children with ADHD. *Psychological Reports*, 105(2):365–371.

Kyungil Kim, Seongjik Lee, and Changhwan Lee. 2015. College students with ADHD traits and their language styles. *Journal of Attention Disorders*, 19(8):687–693.

Dijana Kosmajac. 2020. *Author and Language Profiling of Short Texts*. Ph.D. thesis, Dalhousie University.

Majse Lind, Carla Sharp, and William L Dunlop. 2022. Why, how, and when to integrate narrative identity within dimensional approaches to personality disorders. *J. Pers. Disord.*, 36(4):377–398.

D. P. McAdams. 2013. The psychological self as actor, agent, and author. *Perspectives on Psychological Science*, 8(3):272–295.

Dan P McAdams and Kate C McLean. 2013. Narrative identity. *Curr. Dir. Psychol. Sci.*, 22(3):233–238.

Dan P. McAdams, Jeffrey Reynolds, Martha Lewis, Allison H. Patten, and Phillip J. Bowman. 2001. When bad things turn good and good things turn bad: Sequences of redemption and contamination in life narrative and their relation to psychosocial adaptation in midlife adults and in students. *Personality and Social Psychology Bulletin*, 27(4):474–485.

Kate C McLean, Monisha Pasupathi, and Jennifer L Pals. 2007. Selves creating stories creating selves: a process model of self-development. *Pers. Soc. Psychol. Rev.*, 11(3):262–278.

Matthew B. Miles and A. Michael Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*. Sage Publications.

Mami Miyasaka, Shogo Kajimura, and Michio Nomura. 2018. Biases in understanding attention deficit hyperactivity disorder and autism spectrum disorder in japan. *Front. Psychol.*, 9.

Kathyrn L Mueller and J. Bruce Tomblin. 2012. Examining the comorbidity of language disorders and ADHD. *Topics in language disorders*, 32(3):228–246.

James W Pennebaker and Laura A King. 1999. Linguistic styles: Language use as an individual difference. *J. Pers. Soc. Psychol.*, 77(6):1296–1312.

James W Pennebaker, Matthias R Mehl, and Kate G Niederhoffer. 2003. Psychological aspects of natural language. use: our words, our selves. *Annu. Rev. Psychol.*, 54(1):547–577.

Elena Poznyak, Jessica Lee Samson, Juan Barrios, Halima Rafi, Roland Hasler, Nader Perroud, and Martin Debbané. 2023. Mentalizing in adolescents and young adults with attention deficit hyperactivity disorder: Associations with age and attention problems. *Psychopathology*, pages 1–11.

Halima Rafi, Farnaz Delavari, Nader Perroud, Mélodie Derome, and Martin Debbané. 2023. The continuum of attention dysfunction: Evidence from dynamic functional network connectivity analysis in neurotypical adolescents. *PLoS One*, 18(1):e0279260.

Steve Rathje, Dan-Mircea Mirea, Ilia Sucholutsky, Raja Marjieh, Claire Robertson, and Jay J Van Bavel. 2023. GPT is an effective tool for multilingual psychological text analysis.

Anna Maria Re, Martina Pedron, and Cesare Cornoldi. 2007. Expressive writing difficulties in children described as exhibiting adhd symptoms. *Journal of Learning Disabilities*, 40(3):244–255. PMID: 17518216.

Nina Petersen Reed, Staffan Josephsson, and Sissel Alsaker. 2020. A narrative study of mental health recovery: exploring unique, open-ended and collective processes. *Int. J. Qual. Stud. Health Well-being.*, 15(1):1747252.

Virginia Hill Rice, editor. 2012. *Handbook of stress, coping, and health*, 2 edition. SAGE Publications, Thousand Oaks, CA.

Nanda Rommelse, Marieke van der Kruijs, Jochem Damhuis, Ineke Hoek, Stijn Smeets, Kevin M. Antshel, Lianne Hoogeveen, and Stephen V. Faraone. 2016. An evidenced-based perspective on the validity of attention-deficit/hyperactivity disorder in the context of high intelligence. *Neuroscience & Biobehavioral Reviews*, 71:21–47.

Nader Salari, Hooman Ghasemi, Nasrin Abdoli, Adibeh Rahmani, Mohammad Hossain Shiri, Amir Hossein Hashemian, Hakimeh Akbari, and

Masoud Mohammadi. 2023. The global prevalence of ADHD in children and adolescents: a systematic review and meta-analysis. *Italian Journal of Pediatrics*, 49(1):48.

Jefferson A. Singer, Pavel Blagov, Meredith Berry, and Kathryn M Oost. 2012. Self-defining memories, scripts, and the life story: narrative identity in personality and psychotherapy. *J. Pers.*, 81(6):569–582.

Jefferson A. Singer and Pavel S. Blagov. 2001. *Classification System & Scoring Manual for Self-Defining Memories*. Connecticut College, New London, CT.

Jefferson A Singer and Pavel S Blagov. 2004. The integrative function of narrative processing: Autobiographical memory, self-defining memory, and the life story of identity. In *Studies in self and identity.*, pages 117–138. Psychology Press.

Jefferson A. Singer, Blerim Rexhaj, and Jenna L. Baddeley. 2007. Older, wiser, and happier? comparing older adults' and college students' self-defining memories. *Memory*, 15(8):886–898.

Ana-María Soler-Gutiérrez, Juan-Carlos Pérez-González, and Julia Mayas. 2023. Evidence of emotion dysregulation as a core symptom of adult ADHD: A systematic review. *PLoS One*, 18(1):e0280131.

K Suresh Kumar, AS Radha Mani, T Ananth Kumar, Ahmad Jalili, Mehdi Gheisari, Yasir Malik, Hsing-Chung Chen, and Ata Jahangir Moshayedi. 2024. Sentiment analysis of short texts using svms and vsms-based multiclass semantic classification. *Applied Artificial Intelligence*, 38(1):2321555.

Yla R Tausczik and James W Pennebaker. 2009. The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.*, 29(1):24–54.

Avril Thorne and Kate C. McLean. 2001. Manual for coding events in self-defining memories. Unpublished manuscript.

Avril Thorne, Kate C. McLean, and Amy M. Lawrence. 2004. When remembering is not enough: Reflecting on self-defining memories in late adolescence. *Journal of Personality*, 72(3):513–542.

Sylvie Tordjman, Jacques-Henri Guignard, Carolina Seligmann, Emilie Vanroye, Gregory Nevoux, Jacqueline Fagard, Andrei Gorea, Pascal Mamassian, Patrick Cavanagh, and Sandra Lebreton. 2007. Diagnosis of hyperactivity disorder in gifted children depends on observational sources. *Gifted Talent. Int.*, 22(2):62–67.

Louise Vanden Poel and Dirk Hermans. 2019. Narrative coherence and identity: Associations with psychological well-being and internalizing symptoms. *Front. Psychol.*, 10:1171.

Inna Vogel and Meghana Meghana. 2021. Profiling hate speech spreaders on twitter: SVM vs. Bi-LSTM. In *CLEF 2021– Conference and Labs of the Evaluation Forum*, pages 2193–2200, Bucharest, Romania.

Theodore E A Waters and Robyn Fivush. 2015. Relations between narrative coherence, identity, and psychological well-being in emerging adulthood. *J. Pers.*, 83(4):441–451.

Mark L Wolraich, Joseph F Hagan, Jr, Carla Allan, Eugenia Chan, Dale Davison, Marian Earls, Steven W Evans, Susan K Flinn, Tanya Froehlich, Jennifer Frost, Joseph R Holbrook, Christoph Ulrich Lehmann, Herschel Robert Lessin, Kymika Okechukwu, Karen L Pierce, Jonathan D Winner, William Zurhellen, Subcommittee on Children, and Adolescents with Attention-Deficit/Hyperactive Disorder. 2019. Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics*, 144(4):e20192528.

World Health Organization. 2018. *International Classification of Diseases for Mortality and Morbidity Statistics*, 11th ed. edition. World Health Organization.

Paul J. Yoder. 2006. Predicting lexical density growth rate in young children with autism spectrum disorders. *American Journal of Speech-Language Pathology*, 15(4):378–388.

Susan Young. 2015. *ADHD Child Evaluation (ACE), A diagnostic interview of ADHD in children*. Psychology Services Limited, London.