A Tale of Two Food Adventurers: The Challenges and Triumphs of Repeated Food Exposures in Avoidant/Restrictive Food Intake Disorder

Anonymous Author(s)
Affiliation
Address
email

Abstract

Avoidant/Restrictive Food Intake Disorder (ARFID), a new diagnosis in the DSM-5, is an eating disorder that can emerge in early childhood, threatens optimal physical growth and social-emotional development, and has been reported to persist, for some, well into adolescence or adulthood. Food selectivity more broadly has been reported to be more elevated in families of lower income, while the accessibility and affordability of treatment for mental health patients in the underrepresented group are limited. Therefore, it is crucial to develop accessible, affordable, and effective therapies. We designed a unique clinical study that can be implemented at home, which provides patients with a framework to work towards overcoming the challenges associated with ARFID. During the intervention, participants are filmed and relevant facial information is collected, automatically analyzed with machine learning and computer vision, and delivered to medical experts to enhance the knowledge they use for clinical judgment. We automatically extract affect-related features right after the participants taste or smell a food they labeled as moderately challenging. We observed that facial action units activation provides interesting patterns helpful in understanding the patient's experience throughout the food exposure treatment. This rich information enables quantification of the effectiveness of the currently investigated treatments and differentiation of patient-specific responses to them, potentially leading to scalable personalized medicine for ARFID.

1 Introduction

Parents are in urgent need of strategies that will help their children with clinically severe food avoidance to approach, consume, and enjoy food. As a case in point, Avoidant/Restrictive Food Intake Disorder (ARFID) is an eating disorder that can emerge in early childhood [1, 2, 3], threatens optimal physical growth and social-emotional development [4], and has been reported to persist, for some, well into adolescence or adulthood [1, 2, 3]. ARFID can have a diverse range of presentations yet, a substantial subset of children with ARFID consumes a limited variety of food to the extent that it impairs functioning [4]. Even in the presence of adequate calories for growth, limited dietary variety is associated with numerous health consequences including stomach pain and related digestive problems [5], constipation [6], decreased bone density [4], anemia [7], scurvy [8], as well as other deleterious outcomes related to insufficient vitamins and/or minerals. Limits to dietary variety also have social consequences since it is challenging for a child to find acceptable foods when they eat outside of the home, behavior patterns that may contribute to social isolation. Finally, there is a significant economic burden to families resulting, in part, from excessive food waste when parents have to repeatedly dispose rejected food. These consequences mentioned above regarding inadequate nutrition may be particularly critical during early childhood, given increasing evidence linking poor nutrition during childhood to decrements in future cognitive functioning [9].

Despite possible devastating outcomes of the disorder, receiving mental health treatment in a therapist's office is rapidly becoming obsolete. COVID-19 forced a transformation in the telehealth landscape as video and telephone-based services replaced in-office therapeutic visits. This transformation created unforeseen benefits such as greater access to care (e.g., reduced need for transportation or time off of work). It also may have created some unforeseen benefits in mental health treatment. The mental health service delivery model rested on the assumption that skills learned within a therapist's office would generalize to the outside, lived experiences of the patient. Instead, when such skills are learned within the environment in which they are taught, the potential for skill practice and strengthening may be much greater. Such issues of accessibilty, affordability, and potentiated therapeutic benefit may be particularly relevant for families with a child with ARFID. Picky eating has been shown to be more severe in families of low-income [10]. The management of ARFID typically requires the implementation of strategies at every meal. Such conditions are rife for the development and assessment of tools that can aid parents in the mealtime management of ARFID delivered in the home.

To address the limitations discussed above, we brought together clinical experts in ARFID and machine learning researchers. We designed a home-based treatment that leverages current knowledge in ARFID therapy while it innovates in quantifying novel, more scalable, and objective participant information via machine learning and computer vision techniques. Participants are encouraged to try new foods during repeated exposures [11, 12], in which they gradually try to overcome different stages related to food acceptance. The process is presented in a gamified fashion to engage young children, where each step is related to"climbing a mountain" each time they succeed or take a new phase, they are rewarded, e.g., with stickers (see Figure 1, additional details are presented in Methods). Participants are recorded during each session of the game. We provide preliminary evidence showing that (a) computer vision facial analysis can be implemented as part of an at-home ARFID treatment, (b) participants engage in this framework and can self-record valuable video information with minimal instructions, and (c) that facial information, in particular, facial action units, can be exploited to assess the reactions and emotional state of young individual during the personal journey of an ARFID treatment.

2 Methods

Recruitment and Participants. The study was approved by the university's institutional review board. Due to the online nature of the study, we had the ability to recruit participants from all around the globe. Our recruitment methods were primarily through digital means, most notably different social media platforms. Participants were recruited through Facebook advertisements to targeted groups of parents of children with food avoidance or picky eating, direct referral from medical providers across our university medical center and the community, school newsletters, and our research lab's website. Additionally, our team participated and recruited at several community events to reach out to our local community, and in particular under-represented participants.

Screening. Once a caregiver expressed interest in participating in the study, we evaluated if the participant met the eligibility criteria. These include parent and child having to be proficient in English, and the child being between five to nine years old. In addition, the child had to have met at least one of the following criteria; a neophobia scale sum score of ≥ 29 [13], was considered underweight, had received a diagnosis of feeding disorder or ARFID, had a feeding tube because of an eating disorder, consumed nutritional supplements to help maintain or gain weight, or has marked psychosocial impairment in avoiding of social eating situations.

Data collection. For accessibility and scalability, all the treatments were done at home with participants' caregivers recording the evolution of the child during the multiple exposures to a variety of food, as discussed before. We name this process the "food adventure" to provide the participants with a positive context in which they can approach new foods as an adventure. We gamified their progress with a board (see Figure 1) in which their steps are presented as steps on an adventure while climbing a mountain [14]. Before starting the treatment, each participant defines the steps in the mountain; these are incremental steps of food exposure actions, e.g., looking, touching, smelling, licking, and biting the food. The difficulty level increases with each step, and the last step consists of either finishing the meal or taking multiple bites. For each trial, participants are given food and have to climb each step of the mountain. The participant concludes a session when they complete the last step of the mountain or decide they can not move forward due to disgust. Participants are also instructed to take surveys on the level of disgust toward the food before and after trials. Participants have 15 chances to complete the mountain for each target food.

Facial Features Extraction. Videos are recorded in vertical (portrait) mode, the caregivers use their phone to record the videos, which are then uploaded to our study database in a secure and encrypted fashion. This makes the study scalable and practical, but poses technical challenges. The recorded videos present significant variations of headpose and illumination conditions. We empirically observed these can produce inaccurate face detection and tracking, as well as noisy facial landmark identification. To mitigate these problems. which push machine learning and computer vision to uncharted territories, we combine OpenFace 2.0 [15] and MediaPipe [16] outcomes. For each model, we compare the detected face bounding box and exclude frames for which the intersection over union (IOU) between the two models outputs is lower than 0.5. OpenFace 2.0 algorithm is used to extract facial action units for the subset of validated frames. Since we are interested in evaluating subjects' facial reactions while they are approaching food, we focus on the action units associated to the regions of the nose and eyes. We empirically observed that landmarks and action units associated with the mouth are noisy while participants are eating or approaching food to their face. In this study, we focus in particular on action unit 9, which is associated with the a nose wrinkler movement. This action unit is a relevant proxy for disgust and repulsion [17].

Keyframes annotation. We are interested in evaluating the spontaneous child responses after approaching food (these events include smelling, touching, licking, and biting the food). Clinical experts classified these actions into four levels of difficulty: (i) approaching the food, (ii) oral contact, (iii) tasting, and (iv) eating. Then, they manually annotated the participants videos, and labeled the keyframes where these events took place.

3 Results

The goal of the present work is to show the feasibility of collecting rich clinical information at home, and measuring using computer vision tools picky eaters emotional journey during repeated food exposures. To this end and for illustration in this report, a licensed clinical expert selected two participants with contrasting progress. This selection was agnostic to the information extracted via computer vision, and was based on the expert clinical judgment and information in the electronic health records available. Figure 1 illustrates the intensity of the noise wrinkler facial action unit (AU 9) for a participant that has positive progress (b) and participant without progress (c). Since spontaneous reactions tend to be localized in time [18], each row represents 50 frames (approximately 2.3 seconds) after one of the keyframes defined above (the participant smelled, licked, or bit the food). Average intensity of action units vary across subjects [19, 20], to account for differences across subjects we represent per-subject normalized intensity.

As mentioned above, Figure 1 (b)-(c) shows distinctive differences between the action unit heatmap for the progress participant and the no progress participant. The progress participant showed more activation of nose wrinkler compared to the no progress participants in the early trials of the food adventure. One of the explanations for such phenomena is that it is a result of trying and struggling hard to adjust to the food exposure as a progress participant was able to reach the final step of the mountain for broccoli with six trials, while a no progress participant was not able to eat apples after 15 trials. Compared to the progress participant, the no progress participant has more activation of the nose wrinkler in a later trial of exposure to foods. This could be an example of how food exposure may not be effective or even increase aversion toward the food. Furthermore, this stresses the need for personalized strategies and objective measurements of progress, and provided by machine learning and computer vision.

Interestingly, in the two cases illustrated here progress was not linked to an easier experience, but rather, to the ability of overcoming initial levels of disgust. See for example how self reported feelings (again, Figure 1) after and before the food exposure (left and right of the heatmap, respectively) show an increase level of disgust towards the end of the food adventure, even though they were able to eat the food and their facial expressions showed a reduced level of disgust.

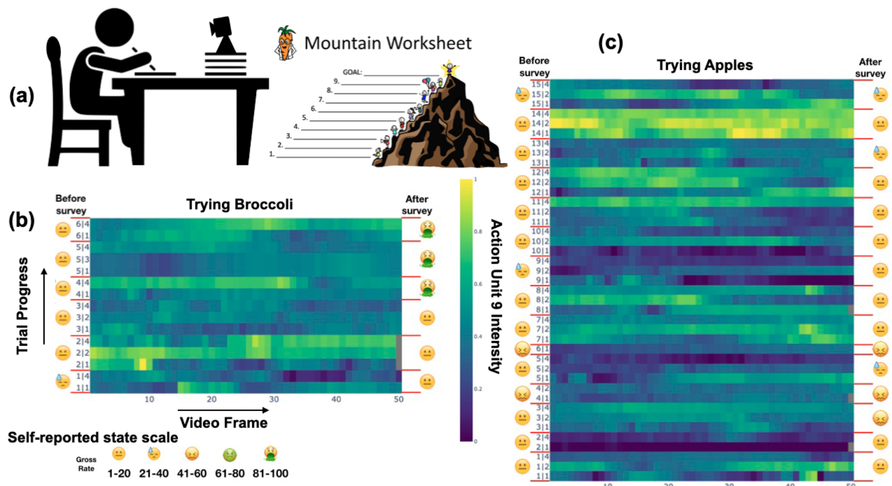

Figure 1: Computer vision based assessment of the "food adventure." (a) Illustrates the participant setup and a the "mountain worksheet" provided to guide the sessions (see Introduction and Methods for details). (b)-(c) Show the changes in normalized intensity of the nose wrinkler for progress (b) and no progress (c) examples. Each row of the heatmap represents a segment of a session, frames are counted on a key-frame (e.g., when the participant tasted or smelled the food). The y-axis (vertical axis) represents number of trial (i.e., how many times they tried this food) and the progress action (i.e., which step of "the mountain" they are attempting). Emojis located on the left/right of the heatmap represent the participants self-reported affect before/after the session.

4 Discussion and conclusions

Computer vision and machine learning hold promise to improve clinical practice by producing scalable, objective, and reliable information. In particular, in the field of pediatric psychiatry, it could alleviate the challenge of measuring reactions and feelings after a food exposure in participants with eating disorders like ARFID. We presented initial evidence that support the feasibility of including computer vision based patient observations at home, and we show that interesting behavioral patterns emerge from the collected data. We observed that assessments based on facial action units might be an accurate alternative to emotional state self report. Children's ability to describe and quantify their internal experience is an area of active and vital research [21]. Discrepant reports between children, their parents, and their healthcare providers on topics as important as pediatric cancer pain point to the growing appreciation for the need to develop tools that accurately depict and communicate a child's experience [22]. The path proposed in the present study has great potential of developing personalized treatment, for example, based on the reactions during the first few food exposures, we could develop data driven and personalized food adventure trajectories.

The present article is framed in a broader study in which we are recruiting over 150 subjects with varying severity of ARFID. Computer vision based tools hold tremendous promise to provide objective clinical information and mitigate disparities in the access and quality of healthcare. Since mobile devices became ubiquitous, the tools discussed in the present work allow clinical experts to access populations traditionally underrepresented both in clinical trials and in access to therapy and healthcare.

Our broader objective is to take steps towards a scalable framework to help young individuals with ARFID; to this end, the community of machine learning need to develop more accurate and robust face/behavioral analysis tools. These need to be able not only to assess a wide range of facial expressions on unconstrained environments, but also to be able to detect when participants are eating, smelling, or licking food. There are tremendous opportunities for the development of body and facial analysis in the contexts of eating, since the problem posses specific and open challenges due to occlusions and facial movement associated with eating. Helping parents to optimize the decisions that they make regarding food purchases, preparation, and presentation will ensure that families can enjoy relaxing, nutritious, and cost-effective meals for generations to come.

References

[1] Sónia Gonçalves, Ana Isabel Vieira, Bárbara César Machado, Renata Costa, Joana Pinheiro, and Eva Conceiçao. Avoidant/restrictive food intake disorder symptoms in children: Associations with child and family variables. *Children's Health Care*, 48(3):301–313, 2019.

[2] Hilde Krom, Liesbeth van der Sluijs Veer, Suzanne van Zundert, Marie-Anne Otten, Marc Benninga, Lotte Haverman, and Angelika Kindermann. Health related quality of life of infants and children with avoidant restrictive food intake disorder. *International Journal of Eating Disorders*, 52(4):410–418, 2019.

[3] Keith E Williams, Helen M Hendy, Douglas G Field, Yekaterina Belousov, Katherine Riegel, and Whitney Harclerode. Implications of avoidant/restrictive food intake disorder (arfid) on children with feeding problems. *Children's Health Care*, 44(4):307–321, 2015.

[4] Debra K Katzman, Mark L Norris, and Nancy Zucker. Avoidant restrictive food intake disorder. *Psychiatric Clinics*, 42(1):45–57, 2019.

[5] Terence M Dovey, Paul A Staples, E Leigh Gibson, and Jason CG Halford. Food neophobia and 'picky/fussy'eating in children: a review. *Appetite*, 50(2-3):181–193, 2008.

[6] D Field, M Garland, and K Williams. Correlates of specific childhood feeding problems. *Journal of Paediatrics and Child Health*, 39(4):299–304, 2003.

[7] F Feillet, A Bocquet, A Briend, J-P Chouraqui, D Darmaun, M-L Frelut, J-P Girardet, D Guimber, Régis Hankard, A Lapillonne, et al. Nutritional risks of arfid (avoidant restrictive food intake disorders) and related behavior. *Archives de Pédiatrie*, 26(7):437–441, 2019.

[8] Megumi Seya, Atsuhiko Handa, Daisuke Hasegawa, Toshihiro Matsui, and Taiki Nozaki. Scurvy: from a selective diet in children with developmental delay. *The Journal of Pediatrics*, 177:331, 2016.

[9] Deborah P Waber, Cyralene P Bryce, Garrett M Fitzmaurice, Miriam L Zichlin, Jill McGaughy, Jonathan M Girard, and Janina R Galler. Neuropsychological outcomes at midlife following moderate to severe malnutrition in infancy. *Neuropsychology*, 28(4):530, 2014.

[10] Katherine R Arlinghaus and Melissa N Laska. Parent feeding practices in the context of food insecurity. *International Journal of Environmental Research and Public Health*, 18(2):366, 2021.

[11] Lean Lipps Birch and Diane Wolfe Marlin. I don't like it; i never tried it: effects of exposure on two-year-old children's food preferences. *Appetite*, 3(4):353–360, 1982.

[12] Jane Wardle, Monica L Herrera, Lucy Cooke, and E Leigh Gibson. Modifying children's food preferences: the effects of exposure and reward on acceptance of an unfamiliar vegetable. *European Journal of Clinical Nutrition*, 57(2):341–348, 2003.

[13] Terence M Dovey, Victoria K Aldridge, Clarissa I Martin, Markus Wilken, and Caroline Meyer. Screening avoidant/restrictive food intake disorder (arfid) in children: Outcomes from utilitarian versus specialist psychometrics. *Eating Behaviors*, 23:162–167, 2016.

[14] Kathryn M Peterson, Cathleen C Piazza, and Valerie M Volkert. A comparison of a modified sequential oral sensory approach to an applied behavior-analytic approach in the treatment of food selectivity in children with autism spectrum disorder. *Journal of Applied Behavior Analysis*, 49(3):485–511, 2016.

[15] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2018)*, pages 59–66. IEEE, 2018.

[16] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Guang Yong, Juhyun Lee, et al. Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*, 2019.

[17] Paul Ekman and Wallace V Friesen. Facial action coding system. *Environmental Psychology & Nonverbal Behavior*, 1978.

[18] Wen-Jing Yan, Qi Wu, Yong-Jin Liu, Su-Jing Wang, and Xiaolan Fu. Casme database: A dataset of spontaneous micro-expressions collected from neutralized faces. In *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, pages 1–7. IEEE, 2013.

[19] Robert Walecki, Vladimir Pavlovic, Björn Schuller, Maja Pantic, et al. Deep structured learning for facial action unit intensity estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3405–3414, 2017.

[20] Angeliki Fydanaki and Zeno Geradts. Evaluating openface: an open-source automatic facial comparison algorithm for forensics. *Forensic Sciences Research*, 3(3):202–209, 2018.

[21] Darren A DeWalt, Heather E Gross, Debbie S Gipson, David T Selewski, Esi Morgan DeWitt, Carlton D Dampier, Pamela S Hinds, I-Chan Huang, David Thissen, James W Varni, et al. Promis® pediatric self-report scales distinguish subgroups of children within and across six common pediatric chronic health conditions. *Quality of Life Research*, 24(9):2195–2208, 2015.

[22] Scott H Maurer, Pamela S Hinds, Bryce B Reeve, Jennifer W Mack, Molly McFatrich, Li Lin, Janice S Withycombe, Shana S Jacobs, Justin N Baker, Sharon M Castellino, et al. Patients, caregivers, and clinicians differ in performance status ratings: Implications for pediatric cancer clinical trials. *Cancer*, 127(19):3664–3670, 2021.