



Disponible en ligne sur
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com



Article de recherche

Amélioration de l'exploration visuelle d'un visage par des enfants avec autisme grâce au ralentissement de la dynamique faciale : une étude préliminaire en oculométrie

Slowing down the flow of facial information enhances facial scanning in children with autism spectrum disorders: A pilot eye tracking study

A. Charrier^a, C. Tardif^a, B. Gepner^{b,*}

^a EA 3273, centre de recherche en psychologie de la connaissance, du langage et de l'émotion (PsyCLÉ), Aix-Marseille université, 13621 Aix-en-Provence, France

^b Laboratoire de neurobiologie des interactions cellulaires et neurophysiopathologie (NICN), UMR CNRS 7259, faculté de médecine Marseille Nord, Aix-Marseille université, 13344 Marseille, France

INFO ARTICLE

Historique de l'article :

Reçu le 9 avril 2014

Accepté le 2 décembre 2015

Disponible sur Internet le xxx

Mots clés :

Autisme
Dynamique faciale
Ralentissement
Fixation visuelle
Oculométrie

RÉSUMÉ

Des particularités dans l'exploration visuelle des visages, notamment des yeux et/ou des lèvres, sont caractéristiques des personnes atteintes de troubles du spectre de l'autisme (TSA) et contribuent à expliquer leurs difficultés de communication sociale et/ou verbale. Elles pourraient être en partie liées à la vitesse excessive des informations véhiculées par les visages lors des interactions. Pour tester cette hypothèse, nous avons enregistré, à l'aide d'un oculomètre, chez 23 enfants avec TSA et 29 enfants typiques, les comportements d'exploration visuelle, i.e. le nombre, le temps total et la durée moyenne des fixations visuelles, sur une scène visuelle montrant le visage d'une narratrice racontant une histoire enfantine dans trois vitesses différentes : temps réel (TR), ralenti à 70% (R70), ralenti à 50% (R50). Nos principaux résultats indiquent que : les enfants avec TSA regardent moins le visage et la bouche de la narratrice, mais autant ses yeux, que les enfants typiques ; les enfants avec TSA font des fixations visuelles d'une durée moyenne significativement plus longue sur la scène visuelle et particulièrement sur la bouche de la narratrice, en condition R50 comparativement à la vitesse TR ; les enfants ayant un autisme léger passent d'autant plus de temps à explorer le visage de la narratrice, et notamment sa bouche et ses yeux, que la vitesse est ralenti. Ces résultats encourageants confirment que la dynamique faciale est trop rapide pour certains enfants autistes et démontrent que le ralentissement de la dynamique faciale peut stabiliser leur regard sur le visage d'autrui et ses traits internes, qui sont cruciaux pour le développement de la communication verbale et socio-émotionnelle.

© L'Encéphale, Paris, 2015.

ABSTRACT

Keywords:
Autism spectrum disorders
Facial dynamics
Slow down
Attention to face
Eye tracking

Objective. – Face and gaze avoidance are among the most characteristic and salient symptoms of autism spectrum disorders (ASD). Studies using eye tracking highlighted early and lifelong ASD-specific abnormalities in attention to face such as decreased attention to internal facial features. These specificities could be partly explained by disorders in the perception and integration of rapid and complex information such as that conveyed by facial movements and more broadly by biological and physical environment. Therefore, we wish to test whether slowing down facial dynamics may improve the way children with ASD attend to a face.

Methods. – We used an eye tracking method to examine gaze patterns of children with ASD aged 3 to 8 ($n = 23$) and TD controls ($n = 29$) while viewing the face of a speaker telling a story. The story was divided into 6 sequences that were randomly displayed at 3 different speeds, i.e. a real-time speed (RT), a slow speed ($S70 = 70\%$ of RT speed), a very slow speed ($S50 = 50\%$ of RT speed). S70 and S50 were displayed

* Auteur correspondant.

Adresse e-mail : [bruno.gepner@univ-amu.fr](mailto;bruno.gepner@univ-amu.fr) (B. Gepner).

thanks to software called Logirai™, aimed at slowing down visual and auditory stimuli simultaneously and without tone distortion. The visual scene was divided into four regions of interest (ROI): eyes region; mouth region; whole face region; outside the face region. The total time, number and mean duration of visual fixations on the whole visual scene and the four ROI were measured between and within the two groups.

Results. – Compared to TD children, children with ASD spent significantly less time attending to the visual scenes and, when they looked at the scene, they spent less time scanning the speaker's face in general and her mouth in particular, and more time looking outside facial area. Within the ASD group mean duration of fixation increased on the whole scene and particularly on the mouth area, in R50 compared to RT. Children with mild autism spent more time looking at the face than the two other groups of ASD children, and spent more time attending to the face and mouth as well as longer mean duration of visual fixation on mouth and eyes, at slow speeds (S50 and/or S70) than at RT one.

Conclusions. – Slowing down facial dynamics enhances looking time on face, and particularly on mouth and/or eyes, in a group of 23 children with ASD and particularly in a small subgroup with mild autism. Given the crucial role of reading the eyes for emotional processing and that of lip-reading for language processing, our present result and other converging ones could pave the way for novel socio-emotional and verbal rehabilitation methods for autistic population. Further studies should investigate whether increased attention to face and particularly eyes and mouth is correlated to emotional/social and/or verbal/language improvements.

© L'Encéphale, Paris, 2015.

1. Introduction

Les troubles du spectre de l'autisme (TSA) appartiennent à la catégorie des troubles neuro-développementaux et sont essentiellement caractérisés par des altérations de la communication et des interactions sociales, et des intérêts restreints et stéréotypés [1]. Parmi ces derniers, on observe très fréquemment des particularités du traitement perceptif des informations sensorielles environnementales. Ainsi, selon les études, 30 % à 100 % des personnes avec TSA présentent, dès leur plus jeune âge, des anomalies perceptives et/ou sensorielles [2], et ce indépendamment de la sévérité de leurs troubles [3]. Parmi ces anomalies, celles liées aux modalités visuelle et auditive peuvent plus spécifiquement contribuer à expliquer leurs troubles de la communication. En effet, pour qu'une communication verbale, émotionnelle et sociale entre des partenaires soit possible, effective et adaptée, chacun des partenaires doit percevoir et intégrer correctement les indices auditifs et visuels échangés lors de l'interaction.

Dans le domaine auditif, la communication verbale requiert des compétences pour discriminer les paroles des interlocuteurs et les différencier des autres sons environnementaux [4]. Or certaines personnes avec autisme ne traitent pas la voix humaine comme étant distincte de sons non vocaux [5] et ont des difficultés à l'utiliser pour réaliser des jugements socio-affectifs [6] ou attribuer des états mentaux à autrui [7]. De plus, des anomalies dans le traitement temporel des sons perturberaient leur perception de la parole et contribueraient au développement atypique du langage dans l'autisme [8,9]. Ce désordre d'intégration temporelle serait d'autant plus marqué que plusieurs canaux sensoriels sont sollicités à la fois. En effet, les personnes avec autisme sont souvent en difficulté pour traiter simultanément des informations issues de plusieurs modalités sensorielles [10], privilégiant un traitement unimodal. Aussi, contrairement à la population générale dont la perception et la compréhension du langage sont améliorées par l'intégration audiovisuelle [11,12], l'inverse est généralement observé dans la population atteinte de TSA [10]. De nombreux témoignages d'adultes avec TSA confirment que regarder quelqu'un tout en l'écoutant est souvent compliqué pour eux.

Dans le domaine visuel, la communication socio-émotionnelle requiert des compétences pour traiter (observer, décoder, comprendre en contexte, etc.) les mouvements biologiques humains (par ex., les mouvements faciaux émotionnels, les mouvements oculaires, les mouvements labiaux, les mouvements et gestes corporels) qui représentent une source d'informations

sociales [13] et contribuent à l'engagement social et à l'inférence des états psychologiques, motivations et intentions d'autrui [14]. Or des troubles de l'identification et de la catégorisation de certains paramètres du mouvement biologique (identification de déplacements, actions ou expressions faciales émotionnelles humaines représentées par des mouvements de points animés) [15,16], mais aussi des troubles perceptifs et perceptivo-moteurs des flux optiques (réactivité oculomotrice à un mouvement de point cohérent, réactivité posturale au mouvement de l'environnement) [17-19] sont fréquents dans cette population. Les enfants et adultes avec TSA présentent souvent une attention atypique (limitée, dispersée, avec des points d'intérêts différents) envers les mouvements biologiques comparativement à la population typique [13]. Ces particularités supposent l'existence de perturbations dans la perception et le traitement perceptivo-moteur ou perceptivo-cognitif du mouvement dès les premières étapes du développement.

Plus spécifiquement, un nombre croissant d'études utilisant l'oculométrie montre des particularités dans les comportements d'exploration visuelle chez les personnes avec autisme, dès les premiers mois de vie. Cette méthode non invasive, ne nécessitant pas de porter un équipement sur la tête ou d'autres dispositifs importuns, permet aux chercheurs de mesurer comment l'observateur, quel que soit son âge, distribue son regard et peut aussi contribuer à répondre à un large éventail de questions scientifiques [20]. En particulier, les personnes avec TSA passent généralement un pourcentage de temps plus faible à regarder les autres individus [21,23] et elles se centrent moins sur les caractéristiques centrales du visage au profit de régions de moindre importance sociale telles que le décor en arrière-plan ou des objets [24,25], et leur balayage visuel est plus erratique, imprécis et désorganisé [26]. Klin et al. [24] ont été les premiers à utiliser des stimuli dynamiques pour mesurer le balayage visuel effectué sur les visages par 15 adolescents avec autisme. Ils ont relevé un temps de fixation moindre sur la région des yeux et deux fois plus important sur les régions de la bouche, mais aussi du corps et des objets, comparativement aux sujets témoins. De surcroît dans cette étude, les régions « bouche » et « objets » se révèlent être des facteurs prédictifs forts de la compétence sociale des sujets avec autisme : un temps important passé à regarder les objets est associé à une adaptation sociale plus faible et un déficit social plus important et, inversement, le temps de fixation sur la région de la bouche est associé à une adaptation sociale plus grande et un déficit social plus faible. Il a été observé dans une autre étude [27] que lors de la présentation de courtes scènes extraites

d'un programme ou d'un film, les enfants avec TSA ont des patterns de regard plus hétérogènes, sont moins attentifs aux visages et aux interactions sociales, détournant leur regard des acteurs pré-maturément pendant les épisodes de parole comparativement aux enfants présentant un trouble spécifique du langage ou un développement typique. Ainsi, les enfants avec TSA auraient des difficultés à observer un visage en mouvement, en particulier lors des échanges de regards et de paroles [27].

Les particularités du comportement visuel et du traitement des informations visuelles et auditives observées chez les personnes avec TSA participent à leurs anomalies communicatives et interactives, selon plusieurs hypothèses contributives, notamment celles de la faiblesse de la cohérence centrale [28] et celle du sur-fonctionnement perceptif [29]. Une autre hypothèse explicative des difficultés d'interaction sociale propose l'existence de désordres précoces du traitement temporel des flux sensoriels, c'est-à-dire un trouble de la perception et de l'intégration des signaux audiovisuels dynamiques, qui s'avéreraient trop rapides et/ou complexes pour être traités en temps réel de manière globale et cohérente [30,31]. Selon cette approche, les enfants avec TSA présenteraient dès leur plus jeune âge une moindre attention pour le visage d'autrui, non pas du fait d'un manque d'intérêt, mais du fait d'une difficulté, voire d'une aversion, à saisir les nombreux traits du visage constamment en mouvement, à les relier simultanément aux stimuli auditifs, et à en extraire un sens global et cohérent avec le contexte. Afin de laisser aux personnes avec TSA une durée suffisante pour traiter simultanément et en temps réel les nombreux signaux visuels et auditifs constamment et rapidement changeants de l'environnement social, Gepner et al. [30,31] ont proposé de ralentir le flux des informations lors des interactions (mouvements du visage et du corps, paroles et langage). Une série de recherches menées auprès d'enfants avec TSA a montré l'effet bénéfique d'une présentation ralentie des stimuli biologiques, visuels et auditifs, qui se traduit par des performances accrues en reconnaissance des expressions faciales émotionnelles et non émotionnelles [32], en imitation faciale et corporelle [33], et par une normalisation de la catégorisation de phonèmes ambigus [34]. Plusieurs études de cas longitudinales suggèrent aussi que le ralentissement des mouvements faciaux et de la parole d'une personne présentée à l'écran d'un ordinateur améliore significativement l'attention, la compréhension verbale et non verbale, la reciprocité sociale et le comportement chez quatre enfants ayant un autisme sévère [35].

Dans la présente étude menée auprès d'enfants avec TSA comparés à des enfants au développement typique, nous souhaitons vérifier l'hypothèse selon laquelle l'exploration visuelle du visage d'une narratrice (racontant une histoire pour enfants) pourrait être améliorée grâce à un ralentissement de la dynamique faciale et vocale de cette narratrice.

Pour ce faire, nous utilisons, d'une part, un logiciel de ralentissement audiovisuel ad hoc (cf. infra) pour ralentir les séquences présentées aux enfants, et un oculomètre pour étudier finement et de manière non invasive les patterns d'exploration visuelle

des enfants avec TSA. Grâce à ce dispositif, nous enregistrons le temps, le nombre et la durée moyenne des fixations visuelles que font les participants sur la scène visuelle. Nous adoptons ainsi les paramètres de mesure fréquemment utilisés dans les études oculométriques [21–26].

Étant donné les résultats d'études menées en oculométrie, montrant des patterns atypiques d'exploration des visages et en particulier de la région des yeux et de la bouche dans la population avec TSA [21–26], nous nous attendons à ce que les enfants avec TSA, comparativement aux enfants au développement typique, fixent moins longtemps les visages, et en particulier les caractéristiques centrales du visage. En outre, en nous appuyant sur l'hypothèse des désordres du traitement temporel des flux multisensoriels [30,31] et les résultats des études antérieures ayant utilisé le ralenti [32–35], nous prédisons que le ralentissement des mouvements faciaux de la narratrice favorisera l'attention visuelle que les enfants avec TSA porteront sur son visage.

2. Méthodologie

Cette étude a reçu l'avis favorable du comité de protection des personnes de Marseille (2011-A00595-36) et a été réalisée après la signature d'un formulaire de consentement éclairé par les parents de tous les participants, et si possible par les enfants eux-mêmes.

2.1. Participants

Le groupe expérimental est composé de 23 enfants, âgés de 3 à 8 ans (âge moyen : 5 ans 8 mois ; DS : 1 an 7 mois), présentant un syndrome d'autisme infantile selon les critères de la CIM-10 [36], ou des troubles du spectre de l'autisme selon ceux du DSM-5. Les enfants ont été recrutés en CAMSP, SESSAD, hôpitaux de jour et auprès de professionnels exerçant en libéral (pédopsychiatres, psychologues, orthophonistes), travaillant dans les Bouches-du-Rhône. L'échelle d'évaluation de l'autisme infantile (Childhood Autism Rating Scale [CARS]) [37] mesurant la sévérité des troubles autistiques a permis de répartir ces vingt-trois enfants en 3 sous-groupes (voir Tableau 1) : un sous-groupe de trois enfants avec autisme léger, ayant un score à la CARS compris entre 30 et 33 ; un sous-groupe de huit enfants avec autisme modéré ayant un score à la CARS compris entre 33 et 38 ; un sous-groupe de 12 enfants avec autisme sévère ayant un score à la CARS compris entre 38 et 45. La présence d'importants troubles du comportement empêchant la participation des enfants aux mesures oculométriques, l'existence de troubles de l'humeur associés, la prise de médicaments affectant leur système oculomoteur, des pathologies oculaires ou auditives sévères (cécité ou surdité congénitale) et des antécédents de troubles neurologiques sévères (épilepsie non équilibrée, encéphalopathie sévère) sont des critères d'exclusion de l'étude. L'âge mental verbal des enfants a été mesuré avec l'échelle de vocabulaire en images Peabody (EVIP) [38] et leurs niveaux de développement en communication et en socialisation ont été mesurés avec l'échelle de comportement socio-adaptatif de Vineland (Vineland

Tableau 1

Caractéristiques des enfants avec autisme et des enfants au développement typique.

	Groupes avec autisme			Groupe témoin M (DS) n = 29
	Sévère M (DS) n = 12	Modéré M (DS) n = 8	Léger M (DS) n = 3	
Sévérité autisme (CARS)	40,25 (1,64)	35,63 (1,23)	30,33 (0,58)	
Âge chronologique (en mois)	63,92 (18,64)	67,63 (15,43)	87,33 (21,94)	71,09 (17,73)
Âge mental verbal (en mois) (EVIP)	11,63 (12,43)	27,75 (13,48)	96,33 (39,63)	
Âge développemental en communication (en mois) (PEP3)	19,36 (4,52)	27,43 (5,10)	66,33 (12,44)	
Âge équivalent en communication (en mois) (VABS)	14,83 (5,29)	21,75 (4,83)	61,67 (31,72)	
Âge équivalent en socialisation (en mois) (VABS)	16,92 (8,85)	23,88 (6,40)	48 (19,67)	

Tableau 2

Histoire des trois petits cochons et découpage en 6 scènes.

Scènes	Texte	Durée (en s)		
		TR	R70	R50
1	Trois petits cochons, Nif-Nif, Naf-Naf et Nouf-Nouf, construisent leur maison. Le premier petit cochon Nif-Nif construit sa maison en paille. Naf-Naf, le deuxième petit cochon fabrique sa maison en bois, et enfin, Nouf-Nouf, le troisième petit cochon bâtit sa maison en pierre	16	23	33
2	Pendant que les 3 petits cochons sont dans leur maison, un loup affamé s'approche de la maison de paille de Nif-Nif. Le loup frappe à la porte « toc, toc, toc » et dit : « Petit cochon, je peux rentrer ? » « Non, non, non » répond Nif-Nif qui tremble de peur dans sa maison de paille	17	25	35
3	« Alors, je vais souffler et ta maison va s'envoler » dit le loup en soufflant très fort « pfuuuuuuuuu » La maison de paille s'envole. Nif-Nif court chez son frère Naf-Naf dans la maison de bois Le loup frappe à la porte : « toc, toc, toc » et dit : « Petits cochons, laissez-moi rentrer »	17	25	35
4	Les 2 petits cochons, tout tremblant, répondent « Non, non, non » Le loup leur dit « Alors, je vais souffler très fort et la maison va s'envoler », et il se met à souffler	17	25	35
5	La maison de bois s'envole et le loup bondit sur les deux petits cochons qui partent en courant vers la maison en pierre de Nouf-Nouf Le loup énervé, frappe à la porte de la maison de Nouf-Nouf : « Gentils petits cochons, je peux entrer ? » « Non », répondent les 3 petits cochons. « Alors, je vais souffler et votre maison s'envolera ! ». Il gonfle ses joues, souffle de toutes ses forces, mais la maison en pierre ne s'envole pas	17	25	35
6	Très en colère le loup monte sur le toit pour rentrer par la cheminée Mais Nouf-Nouf, avec ses 2 frères, allume un feu dans la cheminée Le loup arrive tombe dans le feu, se brûle les fesses et part en courant de douleur : « Aïe, aïe, aïe » Alors, les 3 petits cochons très joyeux font une super fête !	16	23	33

adaptative behavior scale, VABS) [39] auprès de leurs parents et par une observation directe avec le profil psycho-éducatif (PEP 3) [40].

Vingt-neuf enfants témoins au développement typique ont été recrutés auprès d'enseignants dans des établissements scolaires. Ils étaient appariés aux enfants avec autismes sur l'âge chronologique (moyenne : 5 ans et 9 mois ; DS : 1 an et 4 mois) et sur le genre (cf. Tableau 1). Ces enfants témoins étaient dépourvus de troubles psychiatriques, neurologiques et de retards dans les apprentissages, leur niveau scolaire était moyen.

2.2. Matériel

Un oculomètre Tobii T120 Eye Tracker® (Tobii Technology, Stockholm, Suède), composé d'un écran TFT de 17 pouces, est utilisé pour enregistrer les mouvements et fixations oculaires des enfants. Cet oculomètre portable, ressemblant à un écran d'ordinateur ou de télévision, non invasif et disposant d'une grande tolérance aux mouvements de la tête, est donc particulièrement adapté aux enfants en général, et aux enfants autistes en particulier.

Le système Tobii T120 est un système binoculaire qui repose sur l'illumination de la pupille par un rayon lumineux infrarouge, la réflexion cornéenne de ce rayon, et l'enregistrement conjoint de l'angle de réflexion de ce rayon et de la position de la pupille grâce à la vidéo.

La résolution des vidéos est de 1024×764 pixels. La précision du suivi de la position des yeux est de $0,5^\circ$ pour une résolution spatiale de $0,2^\circ$. La fréquence d'enregistrement est de 120 Hz (soit 120 données sont enregistrées par seconde pour chaque œil). Les données sont identifiées par un horodatage et des coordonnées X, Y. L'oculomètre est relié à un ordinateur portable sur lequel est installé le logiciel Studio 3.0 et deux enceintes (LOGITECH X140, puissance 5 Watts) pour amplifier le son des vidéos de format avi. Le logiciel Studio permet de traiter les données brutes enregistrées par l'oculomètre et d'identifier, par l'intermédiaire d'un filtre de fixations (ClearView fixation filter), les fixations et les saccades. Une fixation est considérée comme la période où la position de l'œil dure au minimum 100 ms et ne varie pas plus de $1,2^\circ$ d'angle visuel.

2.3. Stimuli

Les stimuli présentés à l'enfant consistent en des séquences filmées, en couleur, de format avi, dans lesquelles une narratrice

raconte une version abrégée de l'histoire « Les trois petits cochons », avec des variations de la tonalité des phrases et des expressions faciales (cf. Tableau 2). Seul le visage de la narratrice apparaît à l'écran sur un fond neutre gris. L'histoire est décomposée en 6 scènes (S1, S2, S3, S4, S5, S6), durant environ 16,7 s chacune. Chaque scène est présentée dans 3 vitesses différentes :

- une vitesse dite « temps réel » (TR) ;
- une vitesse dite « lente », correspondant à 70 % de la vitesse temps réel (R70) ;
- une vitesse dite « très lente », correspondant à 50 % de la vitesse temps réel (R50).

Les durées des scènes varient en fonction de la vitesse de présentation : en vitesse TR, elles durent en moyenne 16,7 s ; en vitesse R70, elles durent en moyenne 24,3 s ; en vitesse R50, elles durent en moyenne 34,3 s.

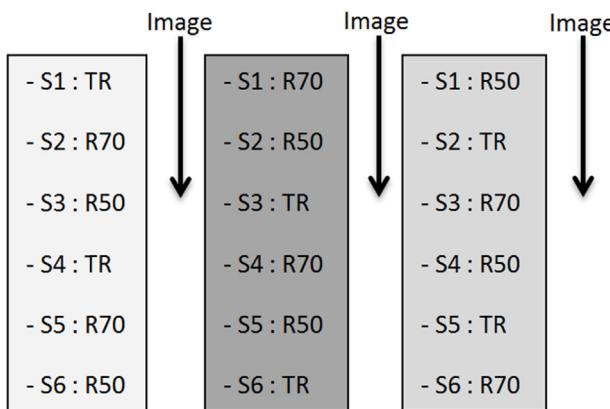
Les vitesses ralenties testées dans cette étude (R70 et R50) sont obtenues grâce à une application numérique spécifiquement développée pour ces recherches, Logirai™ [41], qui permet de ralentir l'image et le son des films de manière parfaitement synchronisée, sans altération de la hauteur de la voix.

L'histoire « Les trois petits cochons », telle qu'elle est présentée aux enfants, est constituée dans un ordre aléatoire de 2 séquences en vitesse temps réel (TR), de 2 séquences en vitesse lente (R70), et de 2 séquences en vitesse très lente (R50).

2.4. Procédure expérimentale

Les enfants avec TSA ont été testés individuellement dans un box expérimental¹, et les enfants typiques ont été testés individuellement dans une salle de taille équivalente à celle du box expérimental et spécialement aménagée à cet effet dans leur école. Le box expérimental et la salle expérimentale à l'école étaient des pièces aveugles (sans fenêtre), aménagés avec le même mobilier (deux tables de 160×80 cm, deux chaises standards et une chaise à dossier de taille ajustable) et les affiches ont été retirées des murs. Le dispositif de mesure (oculomètre, ordinateur, amplificateurs) est le même et positionné de la même façon dans les deux lieux. Afin de contrôler la luminosité dans les deux salles d'expérimentation,

¹ Au laboratoire Parole Langage, UMR CNRS 7309, Aix-Marseille université.

**Fig. 1.** Design expérimental.

nous avons utilisé une source lumineuse incandescente fournissant une lumière constante de 10 lux (contrôlé par un luxmètre). Cette luminosité permet de s'assurer que la pupille des participants est suffisamment large pour éviter les erreurs de calibration et d'enregistrements [42].

L'enfant est placé à environ 60 cm de l'écran de l'oculomètre, correspondant à un angle visuel d'approximativement 30°. Après une phase de calibration, on lui présente l'histoire des trois petits cochons une première fois : la séquence 1 (S1) est présentée en vitesse temps réel (TR), la séquence 2 (S2) en vitesse ralentie à 70% (R70), la S3 en vitesse ralentie à 50% (R50), la S4 en TR, la S5 en R70, la S6 en R50 (voir Fig. 1). À la suite d'une histoire complète, une image extraite d'un dessin animé apparaît au centre de l'écran pendant 4 secondes ; elle permet de remobiliser l'attention de l'enfant sur la scène visuelle. Puis on lui montre l'histoire une seconde fois, chaque scène étant présentée dans une vitesse différente de la première fois. Puis une nouvelle image apparaît pendant 4 secondes. Puis l'histoire est montrée une troisième fois, chaque scène étant présentée à une vitesse différente des première et seconde fois. Enfin, une dernière image représente le design expérimental. L'histoire des trois petits cochons est présentée au total 3 fois (soit 18 scènes) de façon à ce que chaque enfant ait la possibilité de voir chaque scène dans les 3 vitesses différentes. La durée totale de présentation de l'histoire répétée 3 fois et des 3 images est de 476 s (cf. Fig. 1). Le passage d'une scène à une autre scène ou à un stimulus statique se fait à l'aide d'un clic de souris réalisé par l'expérimentateur. Ceci permet d'aménager des pauses selon les besoins et l'attention de l'enfant.

L'histoire des trois petits cochons, découpée en 6 scènes (S1, S2, S3, S4, S5, S6), est présentée 3 fois, de façon à ce que chaque scène soit potentiellement vue par l'ensemble des enfants dans les 3 modalités différentes de vitesse. Une première fois : la S1 est présentée en vitesse temps réel (TR), la S2 en vitesse ralentie à 70% (R70), la S3 en vitesse ralentie à 50% (R50), etc. ; puis, une image

apparaît sur l'écran pendant 4 secondes ; l'histoire est ensuite montrée une seconde fois, chaque scène étant présentée dans une vitesse différente de la 1re fois ; une nouvelle image s'affiche ; et l'histoire est montrée une troisième fois, chaque scène étant présentée dans des vitesses différentes de la première et seconde fois, elle est suivie d'une dernière image.

2.5. Mesures

La scène visuelle présentée à l'écran de l'oculomètre a été divisée en quatre régions d'intérêt :

- la région des yeux de la narratrice ($8,8 \times 2,9$ degrés d'angle visuel) ;
- la région de la bouche de la narratrice ($5,3 \times 2,3$ degrés d'angle visuel) ;
- la région du visage plus celle des cheveux, à l'exclusion des régions des yeux et de la bouche ($15,8 \times 20,5$ degrés d'angle visuel pour la plus grande distance) ;
- la région en dehors du visage (30×17 degrés d'angle visuel) (cf. Fig. 2).

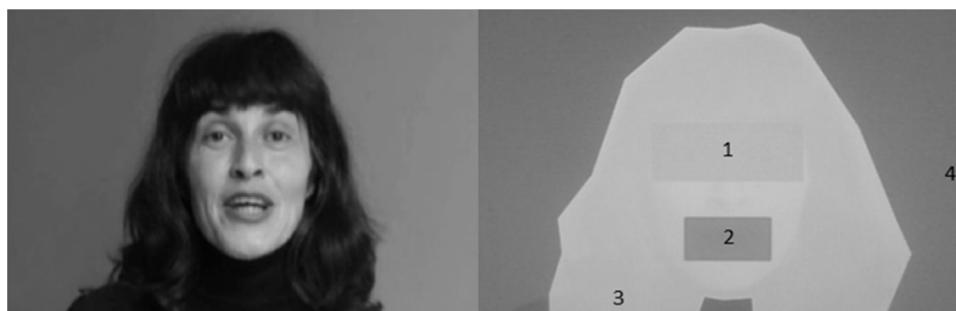
Nous mesurons trois paramètres classiques d'exploration visuelle de la scène visuelle et de chacune des régions d'intérêt : le nombre de fixations, correspondant à la quantité de fixations visuelles faites à l'intérieur d'une région d'intérêt ; le temps total de fixation, qui représente l'addition de l'ensemble des durées de fixations visuelles faites sur une région d'intérêt ; la durée moyenne de fixation visuelle, c'est-à-dire la durée moyenne d'une fixation visuelle faite sur une région d'intérêt.

La scène visuelle est divisée en plusieurs régions d'intérêt :

- la région des yeux ;
- la région de la bouche ;
- la région du visage de la narratrice (comprenant les cheveux et excluant les régions des yeux et de la bouche) ;
- la région en dehors du visage.

2.6. Analyses statistiques

Un test de Shapiro-Wilk a été mené pour vérifier la normalité des variables. L'égalité de variances a été contrôlée par un test de Fischer-Snedecor. Nous avons ensuite utilisé des modèles linéaires généralisés à effet mixte ou GLMM, considérés comme une extension des modèles linéaires généralisés (GLM) [43]. Les GLMM permettent d'étudier le lien entre des variables dépendantes constituées ici par les mesures oculométriques (comprenant le temps total de fixation, la durée moyenne des fixations, et le nombre de fixations visuelles sur la scène visuelle globale et chaque région d'intérêt), et des variables indépendantes dites « explicatives », dans notre cas la vitesse de présentation, tout en

**Fig. 2.** Régions d'intérêt de la scène visuelle présentée aux participants.

prenant en compte la variabilité intra-individuelle. Ces analyses permettent de sélectionner le modèle statistique qui explique le mieux les données de l'étude grâce à l'indice d'ajustement de complexité (AIC). Une fois le meilleur modèle sélectionné, nous avons utilisé des *generalized estimating equations* (GEE) [44] afin de calculer les coefficients standardisés (β) et un seuil significatif pour nos paramètres étudiés. Cette solution permet de traiter les données dans leur intégralité, c'est-à-dire sans constituer de sous-groupes a priori sur la base des mesures obtenues. Certains résultats (différences entre les 3 sous-groupes d'enfants avec TSA) ont néanmoins été présentés pour une lecture plus fine des résultats après nous être assurés que les modèles généraux étaient pertinents.

3. Résultats

Les taux d'enregistrement de données oculométriques sont en moyenne de 27% ($DS = 21$) du temps total de présentation des scènes visuelles pour le groupe avec TSA et de 82% ($DS = 9$) pour le groupe témoin typique. Toutes les mesures de nombre, de temps et de durée moyenne de fixations ont été pondérées par le taux d'enregistrement pour chaque participant.

Nous présenterons d'abord les résultats obtenus sur les trois paramètres de mesures des données oculométriques (nombre de fixations, temps total de fixation et durée moyenne d'une fixation) chez les enfants avec TSA en comparaison de ceux des enfants au développement typique, en fonction de la région d'intérêt, indépendamment de la vitesse de présentation des scènes. Nous présenterons ensuite les données comparatives obtenues au sein du groupe d'enfants avec TSA en fonction de la vitesse de présentation des scènes et du degré de sévérité d'autisme.

3.1. Résultats des enfants avec TSA et des enfants typiques

3.1.1. Scène visuelle globale

Les enfants avec TSA obtiennent des résultats significativement différents de ceux des enfants typiques sur les trois mesures : les enfants avec TSA fixent significativement moins longtemps les scènes visuelles ($b = 16,77$, $DS = 2,89$, $W = 33,77$, $p < .001$), font significativement moins de fixations visuelles ($b = 14$, $DS = 3$, $W = 21,93$, $p < .001$), et celles-ci sont d'une durée moyenne significativement plus courte ($b = .22$, $DS = .05$, $W = 21,57$, $p < .001$).

3.1.2. Régions d'intérêt

Les enfants avec TSA font significativement plus de fixations sur la région en dehors du visage ($b = 7,1$, $DS = 1,73$, $W = 16,87$, $p < .001$) et fixent significativement plus longtemps cette région ($b = 7,1$, $DS = 1,73$, $W = 16,87$, $p < .001$) que les enfants typiques (voir Fig. 3). La durée moyenne des fixations ne diffère pas significativement entre les deux groupes pour cette région.

A contrario, comparativement aux enfants typiques, les enfants avec TSA font significativement moins de fixations sur la bouche ($b = 27,33$, $DS = 2,86$, $W = 91,24$, $p < .001$) et le visage ($b = 36,08$, $DS = 4,72$, $W = 58,25$, $p < .001$), fixent significativement moins longtemps la bouche ($b = 18,38$, $DS = 2,53$, $W = 52,8$, $p < .001$) et le visage ($b = 50,47$, $DS = 2,83$, $W = 319,09$, $p < .001$) et font des fixations d'une durée moyenne significativement plus courte sur la bouche ($b = .27$, $DS = .05$, $W = 29,46$, $p < .001$) et sur le visage ($b = .64$, $DS = .09$, $W = 46,59$, $p < .001$). En revanche, il n'y a pas de différence significative entre les groupes quant au nombre, au temps ou à la durée moyenne des fixations sur la région des yeux.

3.2. Résultats du groupe d'enfants avec TSA

3.2.1. Résultats pour l'ensemble du groupe

3.2.1.1. Scène visuelle globale. Les enfants avec TSA font un nombre de fixations significativement plus important sur les scènes

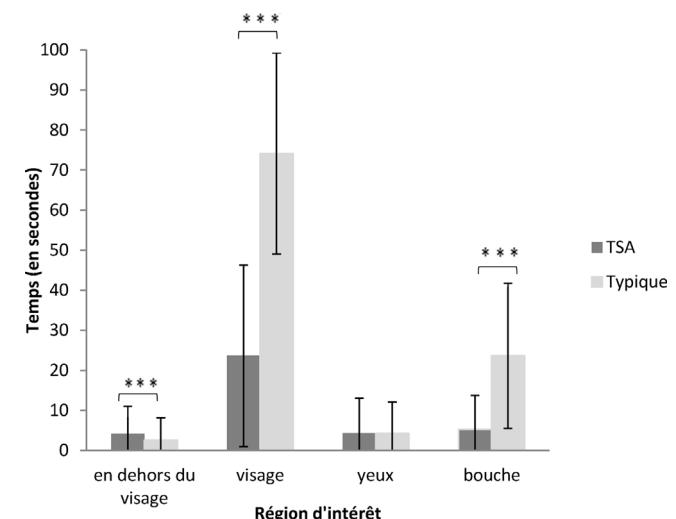


Fig. 3. Temps total de fixation (moyennes et écarts-types) sur les quatre régions d'intérêt (en dehors du visage, visage, yeux, bouche) chez les enfants avec TSA et les enfants témoins au développement typique. *** $p < 0,001$.

présentées en vitesse temps réel comparativement aux scènes en vitesse très lente (R50) ($b = 2,73$, $DS = .96$, $W = 8,13$, $p < .01$). Il n'y a pas de différence significative du nombre de fixations entre les présentations en vitesses lentes (R70) et temps réel (TR), ni entre les vitesses lentes (R70) et très lente (R50) :

- la durée moyenne des fixations est significativement plus longue sur les scènes présentées en vitesse très lente (R50) que sur les scènes en vitesse temps réel (TR) ($b = .04$, $DS = .01$, $W = 7,06$, $p < .01$);
- le nombre et la durée moyenne de fixations ne diffèrent pas significativement entre la vitesse temps réel (TR) et lente (R70), ni entre la vitesse lente et très lente (R70 et R50). Il n'y a pas non plus de différence significative sur le temps total de fixation en fonction de la vitesse de présentation.

Ainsi, bien que le temps total de fixation n'augmente pas significativement en présentation ralentie, l'attention des enfants avec TSA paraît moins dispersée lorsque les scènes sont très lentes, puisque dans cette condition, ils font moins de fixations et chaque fixation dure en moyenne plus longtemps. Autrement dit, les enfants passent moins souvent d'un point de fixation à un autre.

3.2.1.2. Régions d'intérêt. Quelle que soit la région considérée (visage, yeux, bouche ou en dehors du visage), il n'y a pas de différence significative dans le nombre de fixations réalisé et le temps de fixation selon la vitesse de présentation des scènes. En revanche, les fixations ont une durée moyenne plus longue sur la région de la bouche pour les scènes présentées en vitesse très lente (R50) comparativement à celles présentées en vitesse temps réel ($b = .08$, $DS = .03$, $W = 6,80$, $p < .01$) et en vitesse lente ($b = .07$, $DS = .03$, $W = 4,76$, $p = .03$) (Fig. 4). La durée moyenne des fixations réalisées sur cette région ne diffère pas significativement entre les scènes présentées en vitesse lente et temps réel (R70 vs TR).

3.2.2. Comparaisons des sous-groupes d'enfants avec TSA

3.2.2.1. Scène visuelle globale. Les enfants appartenant au sous-groupe avec un autisme léger fixent significativement plus longtemps les scènes que les enfants des deux autres sous-groupes (modéré et sévère) (léger vs modéré : $b = 8,29$, $DS = 4,11$, $W = 4,06$, $p < .05$; léger vs sévère : $b = 9,42$, $DS = 3,38$, $W = 7,76$, $p < .01$) et font des fixations d'une durée moyenne plus longue (léger vs modéré : $b = .25$, $DS = .12$, $W = 4,43$, $p < .05$; léger vs sévère : $b = .27$, $DS = .12$,

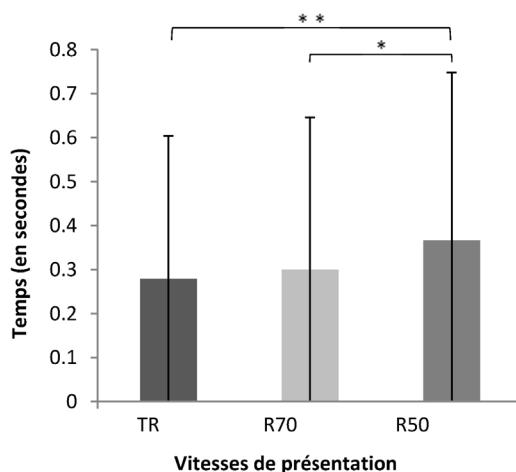


Fig. 4. Durée moyenne de fixation visuelle (moyennes et écarts-types) sur la région de la bouche en fonction de la vitesse de présentation de la scène pour l'ensemble du groupe d'enfants avec TSA. * $p < 0,05$, ** $p < 0,01$; TR: temps réel; R70: ralenti à 70%; R50: ralenti à 50%.

$W = 5,65, p < .05$). En particulier, les enfants du sous-groupe avec un autisme léger fixent significativement plus longtemps les scènes très lentes (R50) que celles présentées en vitesse temps réel (TR) (R50 vs TR : $b = 1,68, DS = .65, W = 6,58, p < .01$). Ils font des fixations d'une durée moyenne significativement plus longue sur les scènes très lentes comparativement aux scènes lentes ($b = .04, DS = .01, W = 15,09, p < .001$) et sur les scènes lentes (R70) et très lentes (R50) comparativement aux scènes en vitesse temps réel (TR) (TR vs R50 : $b = .11, DS = .03, W = 14,62, p < .001$; TR vs R70 : $b = .08, DS = .02, W = 13,34, p < .001$).

Ainsi, les enfants du sous-groupe avec un autisme léger regardent plus longtemps les scènes lorsque celles-ci sont ralenties que lorsqu'elles sont présentées en vitesse TR, et leur durée moyenne de fixation est d'autant plus importante que la vitesse est ralenti.

3.2.2.2. Régions d'intérêt. Les enfants appartenant au sous-groupe avec un autisme léger fixent d'autant plus longtemps la région de la bouche que les scènes sont ralenties (R70 vs R50 : $b = 1,37, DS = .36, W = 14,19, p < .001$; TR vs R50 : $b = 3,55, DS = 1,05, W = 11,41, p < .001$). Les mêmes résultats sont obtenus pour la région du visage (R70 vs R50 : $b = 2,53, DS = 1,19, W = 4,52, p < .05$; TR vs R50 : $b = 3,46, DS = 0,78, W = 19,61, p < .001$) (cf. Fig. 5a).

Les enfants appartenant au sous-groupe avec un autisme léger font également des fixations sur la bouche d'une durée moyenne d'autant plus longue que les scènes sont ralenties (R50 vs TR : $b = .25, DS = .08, W = 10,65, p < .001$; R70 vs TR : $b = .20, DS = .07, W = 6,92, p < .01$). Ils font des fixations sur la région des yeux d'une durée moyenne plus longue pour les scènes très lentes par rapport aux scènes lentes (R50 vs R70 : $b = .08, DS = .03, W = 9,82, p < .01$) et en vitesse temps réel (R50 vs TR : $b = .12, DS = .02, W = 26,72, p < .001$) (cf. Fig. 5b).

Il n'y a pas d'autres différences significatives sur d'autres régions d'intérêt pour ce sous-groupe avec autisme léger. Il n'y a pas non plus de différences significatives pour ces mesures dans les deux autres sous-groupes.

4. Discussion

Cette étude pilote en oculométrie avait pour objectif de mesurer l'impact du ralentissement de stimuli faciaux dynamiques sur la manière dont des enfants avec autisme explorent le visage d'une narratrice racontant une histoire sur un écran d'oculomètre.

Nous avons utilisé un stimulus audiovisuel car des anomalies de la perception visuelle et auditive sont observées dans l'autisme, en particulier au niveau des visages [21–26] et des voix [5–7]. Or les visages comme les voix fournissent des informations sociales pertinentes sur les personnes, et leur traitement simultané est essentiel à la communication et aux interactions sociales humaines. Mais c'est seulement l'impact d'une modification de la vitesse de présentation des stimuli visuels qui a été mesuré dans cette étude.

Nos résultats montrent premièrement que les enfants avec TSA fixent moins longtemps les scènes visuelles comparativement aux enfants ayant un développement typique, et quand ils le font, ils regardent moins le visage et la bouche, et davantage la région en dehors du visage que les enfants typiques. Ces résultats sont en accord avec les études montrant que les enfants avec TSA sont moins attentifs que les enfants typiques aux scènes sociales [21,45] et aux caractéristiques centrales du visage, préférant observer des régions sans importance pour la communication sociale [21–25]. Ainsi, par exemple, Chawarska et al. [22] observent que lors d'un contexte d'échanges dyadiques où des signaux explicites d'engagement social (langage adressé à l'enfant et contacts oculaires) apparaissent sur l'écran de visualisation, les jeunes enfants avec autisme regardent moins la scène de façon générale, notamment le visage de l'orateur et sa bouche, et dirigent plus leur regard vers des régions socialement moins importantes que les enfants témoins au développement typique ou présentant un retard de développement. Par contre, dans des conditions sans contacts oculaires ni paroles, la distribution de l'attention entre les principales caractéristiques de la scène sociale est comparable entre les trois groupes. Ce ne serait pas tant la présentation du visage en soi d'une personne qui perturbe les patterns généraux de regard des enfants avec autisme mais plutôt la présence d'échanges dyadiques [22] et aussi, comme le proposent Gepner et al. [30,31], la présence de stimuli visuels et auditifs rapides et/ou complexes.

Nos résultats confirment également en partie notre hypothèse principale concernant l'impact du ralentissement d'une scène visuelle sur les patterns d'exploration visuelle des enfants avec TSA. En effet, plus la scène visuelle est ralenti, moins les enfants avec TSA font de fixations sur la scène visuelle mais plus la durée moyenne de fixation s'accroît, particulièrement sur la région de la bouche. Ce résultat est particulièrement intéressant puisque si les enfants avec TSA font des fixations plus longues sur la région de la bouche lorsque les scènes sont ralenties, il est très vraisemblable qu'ils puissent bénéficier davantage des informations véhiculées par les mouvements labiaux et de leur congruence avec les sons et phonèmes de la parole, comme l'ont montré Tardif et al. [34] chez certains enfants avec autisme dans des tâches de compréhension de mots phonologiquement proches, et comme l'ont montré Meiss et al. [35] chez des enfants avec autisme sévère, entraînés à observer des scènes audiovisuelles ralenties. Ralentir les informations faciales pourrait ainsi accroître et améliorer leur exploration visuelle du visage en atténuant leur tendance à disperser leur regard et/ou éviter le contact visuel avec le visage d'autrui, notamment ses traits internes. Or, ces informations, en contribuant à la reconnaissance des émotions faciales [44] mais aussi au décodage du langage [12], sont cruciales pour la communication verbale et socio-émotionnelle.

Les analyses comparatives complémentaires réalisées sur les trois sous-groupes d'enfants avec autisme léger, modéré et sévère, montrent que le ralenti bénéficie particulièrement aux enfants ayant des troubles autistiques légers. Les enfants de ce sous-groupe regardent plus longuement la bouche et le visage, et font des fixations d'une durée moyenne plus longue sur la bouche et les yeux que les enfants avec un autisme modéré et sévère. Ils regardent aussi d'autant plus longuement la bouche et le visage, et font des fixations d'une durée moyenne d'autant plus longue sur la bouche et les yeux que la présentation est ralenti. Le ralentissement de

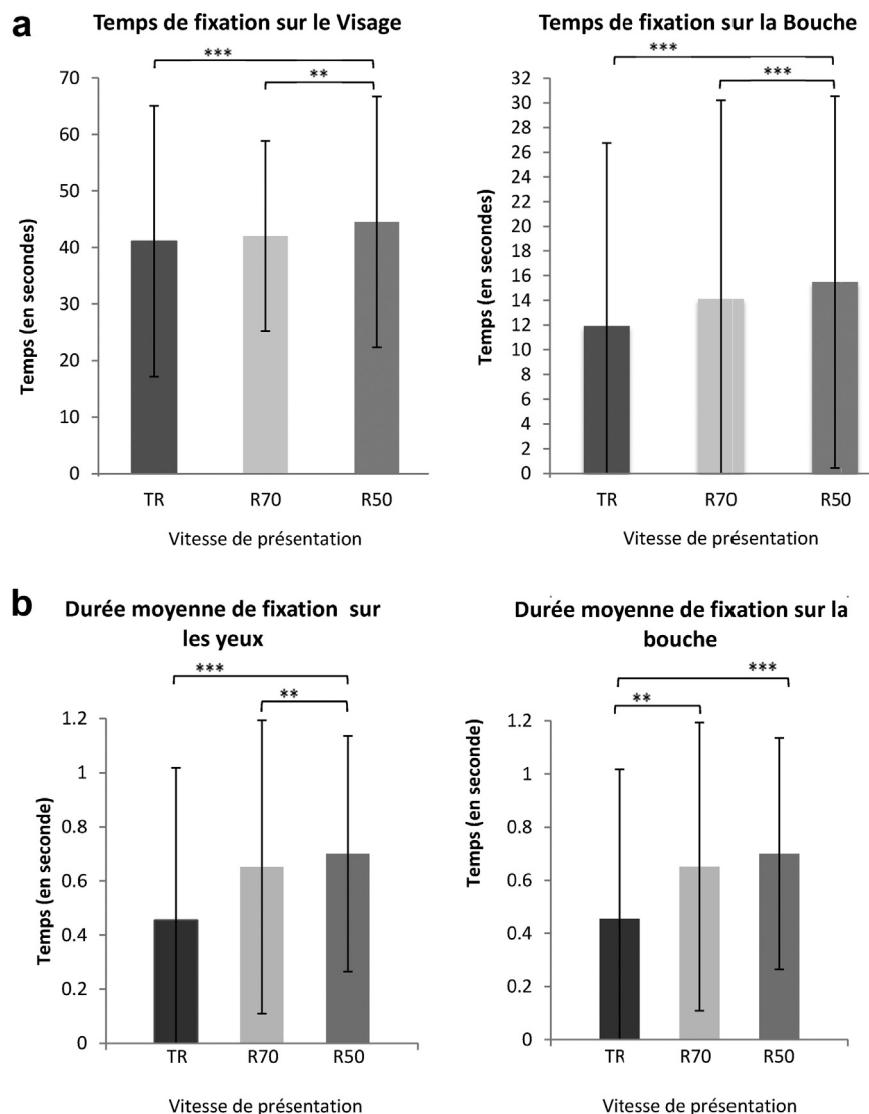


Fig. 5. a : temps total de fixation (moyennes et écarts-types) sur la région du visage et la région de la bouche en fonction de la vitesse de présentation pour le sous-groupe d'enfants avec autisme léger ; b : durée moyenne de fixation (moyennes et écarts-types) sur la région des yeux et la région de la bouche en fonction de la vitesse de présentation dans le sous-groupe d'enfants avec autisme léger. ** $p < 0,01$, *** $p < 0,001$; TR : temps réel ; R70 : ralenti à 70% ; R50 : ralenti à 50%.

la dynamique faciale leur permet d'être plus attentifs aux informations essentielles véhiculées par le visage en contexte social. En stabilisant leur regard sur la région du visage, de la bouche et des yeux, ils ont potentiellement accès aux informations essentielles pour décoder les informations du langage facial, verbal et émotionnel. Ce résultat se distingue de ceux des recherches antérieures montrant des effets bénéfiques du ralenti en particulier pour les enfants ayant les troubles les plus sévères [32,33]. Nous pensons que l'histoire présentée à l'écran est trop compliquée cognitivement, et donc trop peu attractive sur le plan motivationnel, pour les enfants avec autisme modéré et sévère, et reste accessible aux enfants ayant un autisme plus léger et un meilleur niveau de compréhension verbale, explication qu'une étude ultérieure devra s'attacher à démontrer.

Les données de la littérature rapportent que les personnes avec autisme sont en difficulté pour traiter en même temps des informations issues de plusieurs modalités sensorielles, comme regarder quelqu'un tout en l'écouter [46]. Il faudrait vérifier dans une étude ultérieure si l'accroissement du temps et/ou de la durée moyenne de fixations visuelle sur la bouche observés dans le groupe d'enfants avec TSA sont corrélés à une meilleure compréhension de l'histoire racontée, et plus généralement à un meilleur niveau verbal.

Au total, les résultats de notre étude pilote apportent des arguments montrant l'utilité de ralentir la présentation des signaux faciaux pour accroître l'exploration visuelle du visage chez des enfants avec TSA, notamment mais pas seulement ceux ayant un autisme léger, et pour l'orienter vers des indices faciaux cruciaux pour le développement de la communication verbale (lecture labiale) et socio-émotionnelle (lecture des yeux). Des études ultérieures devront chercher à savoir si une attention accrue pour les yeux et les lèvres de visages animés au ralenti s'accompagne ou non d'améliorations développementales, notamment au niveau socio-émotionnel ou verbal, ou comportementales.

Déclaration de liens d'intérêts

Les auteurs déclarent ne pas avoir de liens d'intérêts.

Remerciements

Nous remercions la région Provence-Alpes-Côte-d'Azur pour avoir contribué à financer cette recherche en allouant une bourse à Aurore Charrier, dans le cadre de son doctorat de psychologie (2011–2014). Nous remercions Thomas Arciszewski, et Bruno

Dauvier, respectivement ingénieur d'étude et maître de conférences au centre PsyCLÉ (EA 3273, Aix-Marseille université), ainsi que Nicolas Carvalho, ingénieur de recherche à la Maison de la Recherche (Aix-Marseille université), pour leur aide dans les analyses statistiques. Nous remercions Thierry Legou, ingénieur de recherche au LPL (UMR CNRS 7309), pour son aide technique.

Références

- [1] American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th edition Arlington, VA: American Psychiatric Publishing; 2013.
- [2] Dawson G, Watling R. Interventions to facilitate auditory, visual, and motor integration in autism: a review of the evidence. *J Autism Dev Disord* 2000;30(5):415–21.
- [3] Dunn W, Myles BS, Orr S. Sensory processing issues associated with Asperger syndrome: a preliminary investigation. *Am J Occup Ther* 2002;56:97–102.
- [4] Bergman AS. Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: Bradford Books, MIT Press; 1990.
- [5] Gervais H, Blein P, Boddaert N, et al. Abnormal cortical voice processing in autism. *Nat Neurosci* 2004;7(8):801–2.
- [6] Paul R, Augustyn A, Klin A, et al. Perception and production of prosody by speakers with autism spectrum disorders. *J Autism Dev Disord* 2005;35(2): 205–20.
- [7] Golan O, Baron-Cohen S, Hill JJ, et al. The 'reading the mind in the voice' test-revised: a study of complex emotion recognition in adults with and without autism spectrum conditions. *J Autism Dev Disord* 2007;37:1096–106.
- [8] Oram Cardy JE, Flagg EJ, Roberts W, et al. MEG identifies rapid temporal processing deficit in autism and language impairment. *NeuroReport* 2005;16(4):329–32.
- [9] Roberts TPL, Khan SY, Rey M, et al. MEG detection of delayed auditory evoked responses in autism spectrum disorders: towards an imaging biomarker for autism. *Autism Res* 2010;3(1):8.
- [10] Bogdashina O. Sensory perceptual issues in autism and asperger syndrome: different sensory experiences-different perceptual worlds. London and Philadelphia: JKP; 2003.
- [11] Sommers MS, Tye-Murray N, Spehar B. Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults. *Ear Hear* 2005;26(3):263–75.
- [12] Teinonen T, Aslin RN, Alku P, et al. Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition* 2008;108:850–5.
- [13] Kaiser MD, Pelphrey KA. Disrupted action perception in autism: behavioral evidence, neuroendophenotypes, and diagnostic utility. *Dev Cogn Neurosci* 2012;2:25–35.
- [14] Farroni T, Csibra G, Simion F, et al. Eye contact detection in humans from birth. *Proc Natl Acad Sci U S A* 2002;99:9602–5.
- [15] Blake R, Turner LM, Smoski MJ, et al. Visual recognition of biological motion is impaired in children with autism. *Psychol Sci* 2003;14:151–7.
- [16] Centelles L, Assaiante C, Etchegoyhen K, et al. Compréhension des interactions sociales chez des enfants atteints de troubles du spectre de l'autisme : le langage du corps leur « parle »-t-il ? *Encéphale* 2012;38:232–40.
- [17] Mestre D, Rondan C, Masson G, et al. Évaluation de la vision du mouvement chez des enfants avec autisme au moyen du nystagmus opto-cinétique. *Travaux Interdisciplinaires Lab Parole Langage* 2002;21:192–8.
- [18] Milne E, Swettenham J, Hansen P, et al. High motion coherence thresholds in children with autism. *J Child Psychol Psychiatry* 2002;43(2):255–63.
- [19] Spencer J, O'Brien J, Riggs K, et al. Motion processing in autism: evidence for a dorsal stream deficiency. *NeuroReport* 2000;11(12):2765–7.
- [20] Falck-Ytter T, Bölte S, Gredebäck G. Eye tracking in early autism research. *J Neurodevelopmental Disord* 2013;5(28):1–13.
- [21] Jones W, Carr K, Klin A. Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Arch Gen Psychiatry* 2008;65(8):946–54.
- [22] Chawarska K, Macari S, Shic F. Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biol Psychiatry* 2013;74:195–203.
- [23] de Wit T, Falck-Ytter T, von Hofsten C. Young children with autism spectrum disorder look differently at positive versus negative emotional faces. *Res Autism Spectr Disord* 2008;2:651–9.
- [24] Klin A, Jones W, Schultz R, et al. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch Gen Psychiatry* 2002;59:809–16.
- [25] Shic F, Macari S, Chawarska K. Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. *Biol Psychiatry* 2014;75:231–7.
- [26] Pelphrey KA, Sasson NJ, Reznick S, et al. Visual scanning of faces in autism. *J Autism Dev Disord* 2002;32(4):249–61.
- [27] Hosozawa M, Tanaka K, Shimizu T, et al. How children with specific language impairment view social situations: an eye tracking study. *Pediatrics* 2012;129:1453–60.
- [28] Happé F, Frith U. The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J Autism Dev Disord* 2006;36(1):5–25.
- [29] Mottron L, Dawson M, Soulières I, et al. Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *J Autism Dev Disord* 2006;36(1):27–43.
- [30] Gepner B, Mestre D. Rapid visual-motion integration deficit in autism. *Trends Cogn Sci* 2002;6:455.
- [31] Gepner B, Féron F. Autism: a world changing too fast for a mis-wired brain? *Neurosci Biobehav Rev* 2009;33:1227–42.
- [32] Tardif C, Lainé F, Rodriguez M, et al. Slowing down facial movements and vocal sounds enhances facial expression recognition and facial-vocal imitation in children with autism. *J Autism Dev Disord* 2007;37:1469–84.
- [33] Lainé F, Rauzy S, Tardif C, et al. Slowing down the presentation of facial and body movements enhances imitation performance by children with severe autism. *J Autism Dev Disord* 2011;41:983–96.
- [34] Tardif C, Thomas K, Rey V, et al. Contribution à l'évaluation du système phonologique explicite chez des enfants avec autisme. *Parole* 2002;21:35–72.
- [35] Meiss E, Tardif C, Arciszewski T, et al. Effets positifs d'une exposition à des séquences vidéo ralenties sur l'attention, la communication sociale et les troubles du comportement chez 4 enfants autistes sévères : une étude translationnelle pilote. *Neuropsychiatr Enfance Adolesc* 2015;63:302–9, <http://dx.doi.org/10.1016/j.neurenf.2015.01>.
- [36] World Health Organisation. Mental disorders: a glossary and guide to their classification in accordance with the 11th revision of the international classification of disease (ICD-10). Genève: World Health Organisation; 1996.
- [37] Schopler E, Reichler RJ, DeVellis RF, et al. "Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS)". *J Autism Dev Disord* 1980;10(1):91–103.
- [38] Dunn LM, Thériault-Whalen CM, Dunn LM. Échelle de vocabulaire en images peabody. Adaptation française du Peabody Picture Vocabulary test-revised. Belgique: ATM; 1993.
- [39] Sparrow S, Balla D, Achetti D. Vineland adaptive behavior scale, Aricles Rines. MN: American Guidance Service; 1984.
- [40] Schopler E, Lansing M, Reichler RJ, et al. PEP-3 Profil psycho-éducatif – évaluation psycho-éducative individualisée de la division TEACCH pour enfants présentant des troubles du spectre de l'autisme. Belgique: De Boeck; 2008.
- [41] Tardif C, Gepner B. Logiral [Logiciel PC]; 2012. <http://centrepyscyle-amu.fr/logiral/>.
- [42] Duchowski A. Eye tracking methodology: theory and practice. Springer; 2007.
- [43] Nelder JA, Wedderburn RWM. Generalized linear models. *J Royal Stat Soc* 1972;135(3):370–84.
- [44] Prentice RL, Zhao LP. Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. *Biometrics* 1991;47(3):825–39.
- [45] Nusseck M, Cunningham DW, Wallaven C, et al. The contribution of different facial regions to the recognition of conversational expressions. *J Vis* 2008;8(8):1–23.
- [46] Stevenson RA, Siemann JK, Schneider BC, et al. Multisensory temporal integration in autism spectrum disorders. *J Neurosci* 2014;34(3):691–7.