Disruptors and Co-Designed Interventions for Autistic and ADHD Users in the Audio-Visual Media Context

Craig Cieciura
CVSSP, University of Surrey
Guildford, U.K.
c.cieciura@surrey.ac.uk

Maxine Glancy
BBC Research and Development
MediaCity, Salford, UK
maxine.glancy@bbc.co.uk

Philip J. B. Jackson
CVSSP, University of Surrey
Guildford, U.K.
p.jackson@surrey.ac.uk

IP-based media distribution and machine learning enable increasingly personalised media, yet the accessibility needs of autistic and ADHD communities remain largely underserved. We conducted a multi-stage participatory design study ($N = 8$) with autistic and ADHD participants, combining interviews and co-design activities to identify media disruptors and develop targeted interventions. Through reflexive thematic analysis, we identified key disruptors across sensory, cognitive, and narrative modalities: impeded narrative comprehension, auditory discomfort, dialogue masking in multi-speaker scenes, aversive scene triggers, misaligned pacing, and transient text presentation. From 28 co-designed concepts, we synthesised six functional intervention clusters (spanning intersemiotic translation, audio alteration, content segmentation, and dynamic recaps) and evaluated interactive mockups with a subset of participants ($N = 5$). Evaluation revealed initial evidence of high perceived utility ($Mdn = 3$–5), while highlighting that participants required flexible control to adapt interventions to their personal preferences and viewing contexts. This work provides preliminary design insights for integrating user-tailored accessibility for autistic and ADHD audiences into streaming platforms, laying the groundwork for future generalisable evaluation.

Autism, ADHD, Co-Design, User Study, Audiovisual Media, Accessibility

1. INTRODUCTION

For so long, media has been directed at an audience, designed for the average consumer and produced for a standardised form of presentation. Now, as media technologies and IP-distribution enable more sophisticated enhancements than mere subtitles or captions to address the needs of various individuals and communities, who gets to decide which features are most valuable and deserve engineering into services? This paper represents a people-centred approach to designing technological media adaptations aimed at enhancing the audience experience, combining qualitative participatory investigation alongside our target group with a practical empirical exploration involving emerging technologies.

Audiovisual media plays a central role in satisfying personal and social needs around information, entertainment, and identity formation (Katz et al. 1973; McQuail 1987). While IP-based streaming has expanded access to high-quality content, media accessibility research has predominantly focused on sensory impairments, e.g., blind and low vision (BLV) and deaf hard of hearing (DHH), leaving neurodivergent audiences underserved (Mack et al. 2021; Nevsky et al. 2023). Neurodivergent individuals represent an estimated 15–20% of the population, including 1–3% autistic individuals (Kim et al. 2011; Roman-Urrestarazu et al. 2021; O'Nions et al. 2023) and 3–5% with ADHD (Sarah Stainton 2025; Popit et al. 2024). Autistic viewers often experience sensory processing differences (Seltzer et al. 2003; Hazen et al. 2014; Herrebrøden et al. 2025) alongside emotional recognition challenges and alexithymia (difficulty identifying internal emotional states) (Kinnaird et al. 2019). ADHD involves differences in focus regulation, distractibility (American Psychiatric Association 2022), and working memory variations (Lewis and Brown 2012) that alter media engagement (Solomonidou et al. 2004).

Recent advances in intelligent network computing, e.g., (Porter et al. 2024; Lyko et al. 2024; Liang et al. 2024; Lyko et al. 2023) and machine learning for media manipulation, e.g., (Sardari et al. 2023, 2025; Zhang et al. 2025; Alex et al. 2024; Berghi

© Cieciura et al., Published by BCS Learning and Development Ltd.

et al. 2024; Nadeem et al. 2023; Carion et al. 2025; Team et al. 2025; Shi et al. 2025; Liu et al. 2025) have the potential to enable hyper-specific media adaptations at scale. However, little work leverages these capabilities for autistic and ADHD users.

While autism and ADHD are distinct neurodevelopmental conditions, they exhibit high rates of co-occurrence – estimated at approximately 40% in autistic individuals and 21% in those with ADHD (Rong et al. 2021; Stevens et al. 2016; Hollingdale et al. 2020). Furthermore, prior literature establishes that both populations frequently experience overlapping sensory and cognitive processing differences, particularly regarding sensory over-responsivity, working memory capacity, and executive function. Without conflating these distinct neurotypes, we include both within the scope of our study to explore how dynamic, user-configurable media adaptations might address sensory and cognitive processing demands common to both communities.

Rejecting deficit-based models that treat neurodivergent experiences as impairments to be normalised, we adopt a neurodiversity perspective (Dalton 2013). Assuming mainstream media is designed primarily for neurotypical audiences, autistic and ADHD viewers routinely encounter *disruptors* to their intended media experience. We investigate these disruptors to co-design *interventions* to enable users to personalise their media experience, which we evaluate using experientially functional *mockups*.

Our main research questions are: **RQ1:** What disruptors do autistic and ADHD users experience when engaging with streamed audiovisual media? And **RQ2:** What interventions improve these experiences, and how should they function?

We conducted a qualitative participatory design (PD) study (Simonsen and Robertson 2013) ($N = 8$) involving a pre-task, an artefact-mediated interview, and a co-design workshop. Using inductive, reflexive thematic analysis (Braun and Clarke 2006, 2019) in a graphical environment to analyse artefacts, we developed themes conceptualising the sensory, cognitive, and narrative media disruptors experienced by autistic and ADHD viewers. Finally, we evaluated functional mockups of these interventions with a participant subset. Through this process, this paper makes three primary contributions: (1) the development of themes conceptualising the sensory, cognitive, and narrative media disruptors experienced by autistic and ADHD viewers; (2) a synthesis of six co-designed intervention mechanisms; and (3) a qualitative evaluation of interactive mockups that provides an initial set of practical proposals for streaming platforms for further development and larger scale evaluation.The structure of the paper is as follows: Sections 2 and 3 establish the theoretical context and outline related work in media accessibility. Section 4 details our participatory design methodology. Section 5 presents the developed disruptor themes alongside the co-designed interventions, followed by the mockup evaluations in Section 6. Finally, Section 7 discusses the implications of these findings for future user-tailored streaming features.

2. AUTISM, ADHD, AND MEDIA ACCESSIBILITY

Diagnostic frameworks define autism across two primary dimensions: social communication and interaction differences, and restricted or repetitive patterns of behaviours, interests, or activities, including sensory hyper- and hyposensitivity (Frazier et al. 2012; Marco et al. 2011). Rather than viewing social communication through a deficit model, the *double empathy problem* reframes these difficulties as a bidirectional mismatch in understanding between autistic and non-autistic individuals (Milton 2012). Autistic trait manifestations vary widely across individuals and gender expressions (Lai et al. 2011), accompanied by distinct cognitive strengths and talents (Lampinen et al. 2026; Bal et al. 2022).

ADHD is similarly categorised into two core dimensions: inattention (e.g., difficulty sustaining focus, external distractibility, and disrupted sequential processing) and hyperactivity-impulsivity (e.g., physical restlessness and reduced inhibitory control) (American Psychiatric Association 2022). Diagnosis is divided into three primary presentations: inattentive, hyperactive-impulsive, or combined (Epstein and Loren 2013).

Parsing complex audiovisual media can present significant cognitive load for both autistic viewers (Habib et al. 2019) and ADHDers (Lewis and Brown 2012; Alderson et al. 2010; Sowerby et al. 2011), often linked to visuospatial and phonological working memory variations. Media engagement requires synthesising multimodal audio and visual signs (Chaume 2004; Zabalbeascoa 2008), applying cognitive filtering to prioritise salient details, and integrating these elements into narrative comprehension (Mayer 2005). This synthesis is guided by production “signalling” designed around neurotypical cognitive expectations. When media presentation diverges from a viewer’s cognitive processing style, the intended aesthetic and emotional impact may fail to materialise (Chaume 2004; Romero-Fresco and Chaume 2022). We frame this mismatch as an extension of the double empathy problem (Milton 2012) to audiovisual media – where accessibility barriers stem not from cognitive deficits, but from neurotypical-centric design and production choices.

3. RELATED RESEARCH

Prior media accessibility research predominantly focuses on sensory impairments (Nevsky et al. 2023), including captions for DHH viewers (Neves 2018; Zárate 2021), audio description for BLV audiences (Fryer 2016), and object-based audio customisation (Ward et al. 2018, 2019) (Mack et al. 2021). A smaller body of work explores media engagement among autistic viewers (Judith Fellowes 2011; Lane and Radesky 2019; Hughes et al. 2023) and video accessibility interventions for ADHDers (Zhu et al. 2025; Jiang et al. 2025). Critically, Spiel et al. (2022) argue against the dominant "deficit model" framing in ADHD technology research. A recent related work explores transcreative interventions – adapting narrative delivery rather than literal translation (Romero-Fresco and Chaume 2022) – tailored to the acquired language disorder aphasia (Nevsky et al. 2025).

4. METHOD

To address key critiques of neurodiversity accessibility research, specifically the frequent exclusion of target communities (Mack et al. 2021; Nevsky et al. 2023; Spiel et al. 2022) and deficit-based framing (Milton 2012; Spiel et al. 2022), we adopted a participatory design (PD) approach (Simonsen and Robertson 2013). To safeguard participant well-being, we consulted institutional EDI representatives and an external autistic "expert-by-experience" trainer during protocol development. The study was approved by the institutional Ethics Committee.

4.1. Participants

We recruited eight participants ($N = 8$) through internal mailing lists and posters. Eligibility criteria required participants to be: (1) aged 18+, (2) proficient in English, (3) regular users of video streaming services, and (4) formally diagnosed as autistic, ADHD, or both – including individuals holding one formal diagnosis and currently on an official assessment waiting list for the other. While acknowledging documented structural barriers to clinical diagnosis (Beauchamp et al. 2022; Song et al. 2025; Al-Beltagi et al. 2025; Straub et al. 2022), we restricted inclusion to formally diagnosed or waitlisted individuals to limit sample heterogeneity. Proof of diagnosis was not requested. University-based participants received a £100 gift card upon completion.

Participants were aged 20–52 ($M = 35.0$, $SD = 13.5$) and identified as two women, five men, and one non-binary individual. Diagnostic profiles comprised four autistic participants, one with ADHD, one co-occurring (autistic and ADHD), one autistic participant on an ADHD assessment waitlist, and

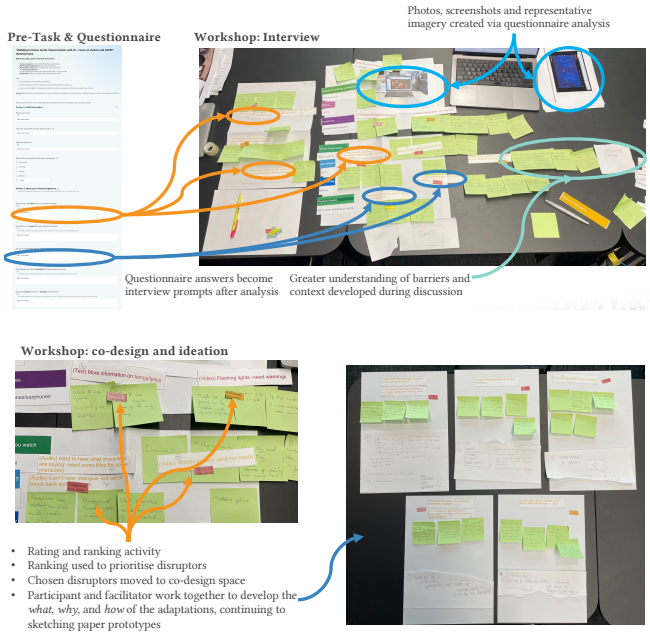


Figure 1: *(a) Questionnaire artefact-mediated workshop interview. (b) Rating, ranking, and intervention co-design workspace. (Images captured during pilot testing using non-participant data.)*

one participant with ADHD on an autism assessment waitlist. Five participants ($N = 5$) from this cohort subsequently returned to complete the qualitative mockup evaluation sessions which were conducted in the following semester. Of those who did not return, two did not respond to emails and the third responded after the sessions had been completed.

4.2. Procedure

The study comprised four sequential stages: an introductory call, a pre-task and questionnaire, a 1:1 interview and co-design workshop, and a final intervention review. The 20-minute online introductory call outlined the study's purpose and stages, addressed practical considerations — including mitigations for discomfort — and allowed participants to ask questions. For the pre-task, participants watched ≥ 15 minutes of self-selected media on BBC iPlayer[1] before completing an online questionnaire. This questionnaire captured viewing context (including participant-uploaded screenshots and environmental photos), general viewing preferences, and perceived disruptors across four core modalities: video, sound, text/graphics, and story/plot.

Workshops (60–90 minutes) were conducted 1:1 with the first or second author, who prepared physical artefacts summarising participant disruptors. Facilitator-participant pairs reviewed, clarified, and annotated these artefacts (Figure 1a). Participants then rated each disruptor's severity on a 5-point

[1] https://www.bbc.co.uk/iplayer

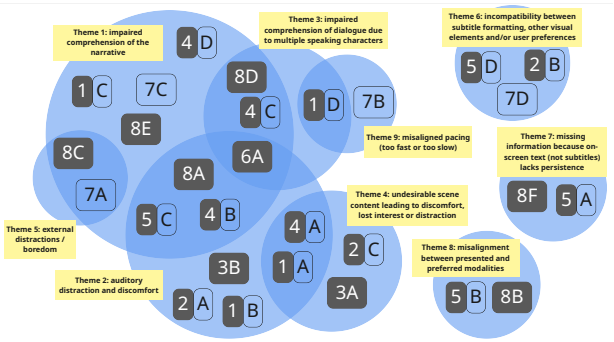


Figure 2: *Overview of disruptor themes. Overlapping regions denote a plurality of experience. Box styling indicates diagnostic status: black fill with white text (autistic), clear fill with black text (ADHD), and split fill (co-occurring diagnosis or active waitlist).*

scale (from *not at all disruptive* to *completely disruptive*) and ranked them by potential for experiential improvement. Finally, participant and facilitator co-designed two to six interventions targeting top-ranked disruptors (Figure 1b), defining each intervention's purpose, targeted media subcomponents, operation, and interaction modalities unconstrained by current technical feasibility.

4.3. Analysis

Interview and co-design workspaces were photographed and digitally synthesised in Miro[2]. We initially annotated each of the 28 co-designed disruptors with summary notes on participant experiences, coping strategies, and solution preferences to support data refamiliarisation. We then conducted an inductive reflexive thematic analysis (Braun and Clarke 2006, 2019). First, analysis focused on the disruptors: artefacts were coded and then spatially clustered into latent disruptor themes, with overlapping positions indicating multidimensional participant experiences (Figure 2). Second, we separately aggregated the ideated interventions into functional clusters based on their primary intended mechanism and affordances (Table 1).

5. RESULTS: MEDIA DISRUPTORS AND CO-DESIGNED INTERVENTIONS

5.1. Understanding Media Disruptors (RQ1)

During workshop interviews, autistic and ADHDer participants described encountering disruptors across multiple sensory and cognitive modalities. In particular, participants reported frequent breakdowns in narrative comprehension resulting from structural plot presentation — such as disconnected scenes, omitted context, and difficulty synthesising causal relationships — which were compounded by primary sensory and cognitive disruptors.

[2]www.miro.com

Table 1: *Primary intervention clusters mapped from content manipulation functions to affordances.*

ID	"Content understanding/ manipulation..."	"...to afford the user..."
1	Scene classification...	...to empower viewing choice
2	Sound replacement or removal...	...to reduce discomfort
3	Sound remixing...	...to reduce distraction
4	Content segmentation...	...to navigate efficiently
5	Identified speaker captioning...	...to follow complex dialogue
6	Dynamic plot summaries...	...to enrich narrative understanding

Auditory disruption was particularly prominent: seven out of eight participants described prior experiences of *auditory distraction and discomfort*. Participants attributed this to specific sound classes, including misophonia triggers (eating and drinking noises, with one participant describing a severe fight-or-flight response), sound effects related to violence or "squelchyness", excessive timbral or high-frequency spectral content, and sudden volume spikes during advertisements.

Participants also reported disruptors relating to text presentation and pacing. Two participants (one diagnosed with ADHD and on the waiting list for an autism assessment, and one autistic) described repeatedly rewinding content to locate transient on-screen text and graphics (e.g., contributor introductions) that vanished before they could fully process them. Playback speed was also highlighted as disruptive for opposing reasons: one participant required slower playback to process dense dialogue due to auditory processing lags (Blomberg et al. 2019), whereas another reported needing accelerated playback to prevent under-stimulation and maintain focus. Finally, four autistic participants described distress resulting from unexpectedly encountering specific aversive scene types (e.g., violence, interpersonal conflict, or scenes lacking context). Because causes of distress were highly specific, and sometimes mediated by whether an occurrence felt narratively motivated, standard content warnings were described as vague or insufficient. Consequently, participants reported experiencing conflict between continuing with content in which they held substantial emotional investment, versus abandoning a series to avoid further distress.

5.2. Co-Designed Intervention Mechanisms (RQ2)

Following the interviews, participants engaged in co-design to generate targeted interventions addressing their disruptors. While concepts addressing standard subtitle or text styling were excluded due to existing coverage in accessibility literature (e.g.,

(Gorman et al. 2021) for personalisable subtitles), the remaining six functional clusters formed the basis for our interactive mockup development:

Scene Classification to Empower [...] Choice. Affords agency over scene triggers without breaking narrative flow, pairing scene-skipping with text summaries or synthesised audio recaps, coupled with placeholder images, preserving plot continuity.

Sound Replacement or Removal. Enables viewers to select, isolate, and eliminate specific distress-causing sound classes (e.g., eating sounds (misophonia) or sounds related to violence) without degrading the remaining audio.

Sound Remixing. Provides relative volume level adjustments across groups of sounds, allowing users to lower background music or effects relative to speech to reduce auditory distractions.

Content Segmentation. Supports re-engagement following attention lapses or processing lags via interactive chapters and segments, with description of content within, interactive transcripts, and context-aware navigation controls (jumping by sentence, speaker change, or segment).

Identified Speaker Captioning. Resolves multi-speaker dialogue parsing through intersemiotic translations (Jakobson 2021), using colour-coded speaker chips, independent speaker-attributed subtitles, and contextual contributor fact-files.

Dynamic Plot Summaries. Supports working memory variations by clarifying narrative causality through current-scene summaries, subplot recaps, character cards, and previous episode summaries.

6. EVALUATION OF INTERVENTION MOCKUPS

We evaluated interactive mockups with five participants ($N = 5$), a subset of our initial workshop cohort ($N = 8$). Interventions were developed iteratively – progressing from workshop co-design sketches and wireframes to functional mockups using content examples suggested during prior workshop sessions. Screenshots of each of the mockups, created using Figma Make[3], appear in Appendix A. For each intervention, participants completed a qualitative walk-through – evaluating both the illustrated example and potential hypothetical applications – and rated their agreement with the statement *"This intervention would be useful in my media watching/listening"* on a 5-point Likert-type scale (1 = Strongly disagree, 5 = Strongly agree; Figure 3).

[3]https://www.figma.com/make/

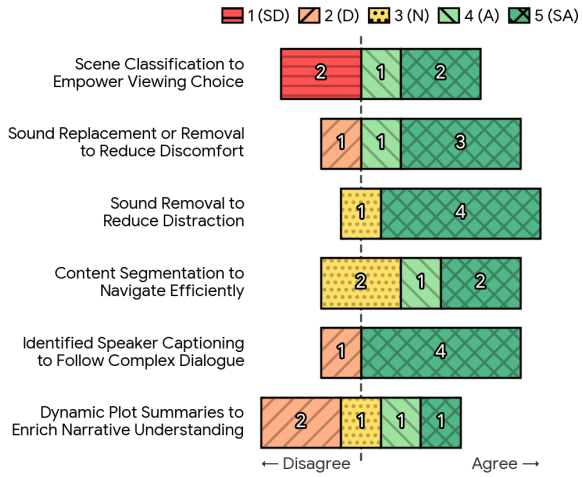


Figure 3: *Participant agreement with the statement "This intervention would be useful in my media watching/listening" ($N = 5$). Segment widths and tone-on-tone hatch patterns reflect answer distributions on a 5-point Likert-type scale (1 = Strongly disagree, 5 = Strongly agree).*

Sound Replacement or Removal ($Mdn = 5$). Evaluating the intervention illustrated with a medical drama example (muting crowd noise and sirens), participants valued eliminating auditory discomfort. However, ID8 observed that muting sirens removed necessary plot context, while ID2 noted that source-separation algorithms introduced phase artefacts potentially causing listening fatigue. Consequently, participants expressed a shared preference for continuous volume adjustment over binary removal (ID2, ID6, ID8). In hypothetical discussions, participants identified key triggers for total removal, such as misophonia from eating noises in films (ID2, ID6), crying infants (ID3), or jarring reality TV sound effects (ID8).

Sound Remixing ($Mdn = 5$). Evaluating the intervention illustrated with a reality show example, participants strongly endorsed adjusting relative audio levels. Lowering background music aided focus and restored speech intelligibility without stripping narrative immersion (ID1, ID3, ID8).

Identified Speaker Captioning ($Mdn = 5$). Evaluating the intervention illustrated with a debate show example, ID6 and ID8 highlighted that colour-coded subtitle tags provided structured visual cues that reduced the cognitive effort of tracking multi-speaker dialogue compared to standard captions. Contextual fact-files provided helpful reference during debate content (ID2, ID8), though ID6 deemed them unnecessary for fictional drama.

Content Segmentation ($Mdn = 4$). Evaluating the intervention illustrated with a documentary example,

ID1, ID3, and ID8 responded positively to interactive segment summaries and transcripts, and context-aware navigation controls, as a tool to recover from attention lapses. Regarding UI layout, ID2 cautioned that dense navigation overlays would be cumbersome on low-power streaming devices, while ID6 and ID8 requested inset or collapsible layouts to avoid obscuring video.

Scene Classification ($Mdn = 4$). Evaluating the intervention illustrated with an action drama example, participants seeking to skip scenes strongly preferred synthesised audio recaps over dense text blocks, which broke viewing immersion (ID2, ID3, ID8). ID3 noted that the audio recap felt almost as graphic as the scene itself, but envisioned high utility for receiving necessary narrative information without having to watch content for its full duration. Conversely, ID6 and ID8 – who do not personally avoid violent content – rated scene-skipping as not useful ($1 =$ Strongly disagree).

Dynamic Plot Summaries ($Mdn = 3$). Evaluating the intervention illustrated with a comedy mockumentary example, participants with working memory variations valued current-scene and previous-scene summaries for clarifying narrative causality (ID1, ID3). In contrast, ID2 and ID6 rated these scene-level summaries lower ($2 =$ Disagree), explaining that explicit summaries could reduce the satisfaction of narrative deduction in fiction. However, both noted potential utility for re-orienting after viewing gaps or tracking dense non-fiction.

7. DISCUSSION AND CONCLUSION

IP-based media delivery and machine learning allow for increasingly personalised media experiences. However, autistic and ADHD communities have historically received limited focused research regarding their specific media access needs and interventions to improve them. We engaged with eight participants with autism or ADHD diagnoses through a multi-stage participatory design process, identifying key disruptors spanning narrative comprehension, sensory discomfort, multi-speaker dialogue parsing, aversive scene content, and pacing misalignments. Evaluating six interactive mockups with five participants demonstrated positive preliminary evidence of perceived utility ($Mdn = 3$–5), showing how inter-semiotic translation and transcreation approaches, across audio, visual, and text modalities, as well as context-aware segmentation and summarisation, can mitigate disruption for those with differing cognitive processing styles. Crucially, evaluation feedback highlighted that feature utility was strongly context-dependent, with participants requiring flexible control to adapt how and when interventions were applied based on personal preference and media genre. While interventions like audio manipulation and synthesised recaps represent forms of *transcreation* (Romero-Fresco and Chaume 2022) that alter original production choices, balancing authorial intent with individual cognitive and sensory needs is essential for advancing a universalistic approach to media accessibility (Greco and Jankowska 2020).

Our findings align with prior research that explored narrower intervention scopes, non-audiovisual contexts, or high-level recommendations. For instance, (Zhu et al. 2025) evaluated prototypes targeting distraction reduction, audio/visual/caption highlighting, and content segmentation for ADHD viewers. Meanwhile, (Jiang et al. 2025) proposed design recommendations for adaptive focus and multisensory delivery in video. In a social media context, (Barros Pena et al. 2023) recommended empowering autistic users to filter distressing content and clarify tone. Concepts from each of these studies are encapsulated across our six co-designed interventions.

Given the formative nature of this qualitative work, our initial sample ($N = 8$ workshops; $N = 5$ evaluations) yields preliminary design insights rather than generalisable findings across broader autistic and ADHD populations. To extend these insights, future research should further evaluate our six interventions, as well as co-design additional concepts, with broader, community-sourced cohorts in real-world streaming environments. This would also help bypass the systemic recruitment barriers noted in our study.

By allowing creatives to preview user-tailored adaptations within production pipelines, media platforms could move beyond static, one-size-fits-all remediation while preserving creative intent. While mainstream media accessibility has historically prioritised sensory interventions, this work provides a preliminary set of co-designed practical proposals for developing flexible streaming features that support the cognitive and narrative needs of autistic and ADHD audiences.

ACKNOWLEDGEMENT

Thanks to UKRI-BBC Prosperity Partnership "AI4ME: Future Personalised Object-Based Media Experiences Delivered at Scale Anywhere" EP/V038087. We also thank Joanna Rusznica for her contribution to the initial design of the study. The authors acknowledge the use of AI for copy-editing and language refinement during the preparation of this manuscript. The human authors reviewed, verified, and edited all generated revisions, and take full responsibility for the content of this paper.

REFERENCES

Al-Beltagi, M., B. S. Mani, E. M. Hantash, A. A. Al Zahrani, and O. Toema (2025, December). Challenges in diagnosing attention-deficit/hyperactivity disorder in pediatric practice: A regional and global perspective. *World J. Clin. Pediatr. 14*(4), 111684.

Alderson, R. M., M. D. Rapport, K. L. Hudec, D. E. Sarver, and M. J. Kofler (2010, May). Competing Core Processes in Attention-Deficit/Hyperactivity Disorder (ADHD): Do Working Memory Deficiencies Underlie Behavioral Inhibition Deficits? *J. Abnorm. Child Psychol. 38*(4), 497–507.

Alex, T., S. Ahmed, A. Mustafa, M. Awais, and P. J. Jackson (2024, April). Max-AST: Combining Convolution, Local and Global Self-Attentions for Audio Event Classification. In *ICASSP 2024 - 2024 IEEE Int. Conf. Acoust. Speech Signal Process. ICASSP*, pp. 1061–1065.

American Psychiatric Association (2022). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5-TR* (5th ed., text rev. ed.). Washington, DC: American Psychiatric Association.

Bal, V. H., E. Wilkinson, and M. Fok (2022, January). Cognitive profiles of children with autism spectrum disorder with parent-reported extraordinary talents and personal strengths. *Autism Int. J. Res. Pract. 26*(1), 62–74.

Barros Pena, B., N. Koteyko, M. Van Driel, A. Delgado, and J. Vines (2023, April). "My Perfect Platform Would Be Telepathy" - Reimagining the Design of Social Media with Autistic Adults. In *Proc. 2023 CHI Conf. Hum. Factors Comput. Syst.*, CHI '23, New York, NY, USA, pp. 1–16. Association for Computing Machinery.

Beauchamp, M. L. H., K. Amorim, S. N. Wunderlich, J. Lai, J. Scorah, and M. Elsabbagh (2022, August). Barriers to access and utilization of healthcare services for minority-language speakers with neurodevelopmental disorders: A scoping review. *Front. Psychiatry 13*.

Berghi, D., C. Cieciura, F. Einabadi, M. Glancy, O. C. Camilleri, P. Foster, A. Nadeem, F. Sardari, J. Zhao, M. Volino, A. Mustafa, P. J. B. Jackson, and A. Hilton (2024, July). ForecasterFlexOBM: A Multi-View Audio-Visual Dataset for Flexible Object-Based Media Production. In *2024 IEEE Int. Conf. Multimed. Expo ICME*, pp. 1–6.

Blomberg, R., H. Danielsson, M. Rudner, G. B. W. Söderlund, and J. Rönnberg (2019, July). Speech Processing Difficulties in Attention Deficit Hyperactivity Disorder. *Front. Psychol. 10*.

Braun, V. and V. Clarke (2006, January). Using thematic analysis in psychology. *Qual. Res. Psychol. 3*(2), 77–101.

Braun, V. and V. Clarke (2019, August). Reflecting on reflexive thematic analysis. *Qual. Res. Sport Exerc. Health 11*(4), 589–597.

Carion, N., L. Gustafson, Y.-T. Hu, S. Debnath, R. Hu, D. Suris, C. Ryali, K. V. Alwala, H. Khedr, A. Huang, J. Lei, T. Ma, B. Guo, A. Kalla, M. Marks, J. Greer, M. Wang, P. Sun, R. Rädle, T. Afouras, E. Mavroudi, K. Xu, T.-H. Wu, Y. Zhou, L. Momeni, R. Hazra, S. Ding, S. Vaze, F. Porcher, F. Li, S. Li, A. Kamath, H. K. Cheng, P. Dollár, N. Ravi, K. Saenko, P. Zhang, and C. Feichtenhofer (2025, November). SAM 3: Segment Anything with Concepts. arXiv [Preprint]. arXiv:2511.16719. Available at: http://arxiv.org/abs/2511.16719 (Accessed: 2026-01-09T21:56:04Z).

Chaume, F. (2004). Film studies and translation studies: Two disciplines at stake in audiovisual translation. *Meta 49*(1), 12–24.

Dalton, N. S. (2013, April). Neurodiversity & HCI. In *CHI 13 Ext. Abstr. Hum. Factors Comput. Syst.*, CHI EA '13, New York, NY, USA, pp. 2295–2304. Association for Computing Machinery.

Epstein, J. N. and R. E. A. Loren (2013, October). Changes in the Definition of ADHD in DSM-5: Subtle but Important. *Neuropsychiatry 3*(5), 455–458.

Frazier, T. W., E. A. Youngstrom, L. Speer, R. Embacher, P. Law, J. Constantino, R. L. Findling, A. Y. Hardan, and C. Eng (2012, January). Validation of Proposed *DSM-5* Criteria for Autism Spectrum Disorder. *J. Am. Acad. Child Adolesc. Psychiatry 51*(1), 28–40.e3.

Fryer, L. (2016, April). *An Introduction to Audio Description: A Practical Guide*. London: Routledge.

Gorman, B. M., M. Crabb, and M. Armstrong (2021, May). Adaptive Subtitles: Preferences and Trade-Offs in Real-Time Media Adaption. In *Proc. 2021 CHI Conf. Hum. Factors Comput. Syst.*, CHI '21, New York, NY, USA, pp. 1–11. Association for Computing Machinery.

Greco, G. M. and A. Jankowska (2020). Media accessibility within and beyond audiovisual translation. In *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*, pp. 57–81. Springer.

Habib, A., L. Harris, F. Pollick, and C. Melville (2019). A meta-analysis of working memory in individuals with autism spectrum disorders. *PloS One 14*(4).

Hazen, E. P., J. L. Stornelli, J. A. O'Rourke, K. Koesterer, and C. J. McDougle (March/April 2014). Sensory Symptoms in Autism Spectrum Disorders. *Harv. Rev. Psychiatry 22*(2), 112.

Herrebrøden, M. K., A. Dechsling, G. Marentakis, and A. Nordahl-Hansen (2025, June). Audio Technology for Autistic Persons with Auditory Sensory Differences—A Scoping Review. *Rev. J. Autism Dev. Disord.*.

Hollingdale, J., E. Woodhouse, S. Young, A. Fridman, and W. Mandy (2020, October). Autistic spectrum disorder symptoms in children and adolescents with attention-deficit/hyperactivity disorder: A meta-analytical review. *Psychol. Med. 50*(13), 2240–2253.

Hughes, L., N. Kargas, M. Wilhelm, H. S. Meyerhoff, and J. Föcker (2023, October). The Impact of Audio-Visual, Visual and Auditory Cues on Multiple Object Tracking Performance in Children with Autism. *Percept. Mot. Skills 130*(5), 2047–2068.

Jakobson, R. (2021). On linguistic aspects of translation. In *The Translation Studies Reader*, pp. 156–161. Routledge.

Jiang, L., W. Ko, S. Yuan, T. Shende, and S. Azenkot (2025, April). Shifting the Focus: Exploring Video Accessibility Strategies and Challenges for People with ADHD. In *Proc. 2025 CHI Conf. Hum. Factors Comput. Syst.*, CHI '25, New York, NY, USA, pp. 1–16. Association for Computing Machinery.

Judith Fellowes (2011, August). Autistic spectrum, captions and audio description.

Katz, E., J. G. Blumler, and M. Gurevitch (1973). Uses and Gratifications Research. *Public Opin. Q. 37*(4), 509–523.

Kim, Y. S., B. L. Leventhal, Y.-J. Koh, E. Fombonne, E. Laska, E.-C. Lim, K.-A. Cheon, S.-J. Kim, Y.-K. Kim, H. Lee, D.-H. Song, and R. R. Grinker (2011, September). Prevalence of Autism Spectrum Disorders in a Total Population Sample. *Am. J. Psychiatry*.

Kinnaird, E., C. Stewart, and K. Tchanturia (2019, January). Investigating alexithymia in autism: A systematic review and meta-analysis. *Eur. Psychiatry 55*, 80–89.

Lai, M.-C., M. V. Lombardo, G. Pasco, A. N. V. Ruigrok, S. J. Wheelwright, S. A. Sadek, B. Chakrabarti, M. A. Consortium, and S. Baron-Cohen (2011, June). A Behavioral Comparison of Male and Female Adults with High Functioning Autism Spectrum Conditions. *PLOS ONE 6*(6), e20835.

Lampinen, L. A., J. Singer, X. Wang, B. VanHook, E. Wilkinson, and V. H. Bal (2026, January). Self-reported strengths and talents of autistic adults. *Autism Int. J. Res. Pract. 30*(1), 37–48.

Lane, R. and J. Radesky (2019, June). Digital Media and Autism Spectrum Disorders: Review of Evidence, Theoretical Concerns, and Opportunities for Intervention. *J. Dev. Behav. Pediatr. JDBP 40*(5), 364–368.

Lewis, D. and V. Brown (2012). Individuals with ADHD and the cognitive processing of multimedia. In *Soc. Inf. Technol. Teach. Educ. Int. Conf.*, pp. 4645–4649. Association for the Advancement of Computing in Education (AACE).

Liang, Y., T. Lyko, M. Nilsson, P. Farrow, S. Appleby, M. Broadbent, and N. Race (2024, June). Playing Catch-Up: Evaluating Playback Speed Control in Low-Latency Live Streaming. In *2024 16th Int. Conf. Qual. Multimed. Exp. QoMEX*, pp. 270–273.

Liu, X., Z. Zhu, H. Liu, Y. Yuan, Q. Huang, M. Cui, J. Liang, Y. Cao, Q. Kong, M. D. Plumbley, and W. Wang (2025). WavJourney: Compositional Audio Creation With Large Language Models. *IEEE Trans. Audio Speech Lang. Process. 33*, 2830–2844.

Lyko, T., M. Broadbent, N. Race, M. Nilsson, P. Farrow, and S. Appleby (2024, February). Improving quality of experience in adaptive low latency live streaming. *Multimed. Tools Appl. 83*(6), 15957–15983.

Lyko, T., Y. Elkhatib, R. Ramdhany, and N. Race (2023, June). Differential QoE in Picture-in-Picture Gaming Videos: A Subjective Study. pp. 221–223.

Mack, K., E. McDonnell, D. Jain, L. Lu Wang, J. E. Froehlich, and L. Findlater (2021, May). What Do We Mean by "Accessibility Research"? A Literature Survey of Accessibility Papers in CHI and ASSETS from 1994 to 2019. In *Proc. 2021 CHI Conf. Hum. Factors Comput. Syst.*, CHI '21, New York, NY, USA, pp. 1–18. Association for Computing Machinery.

Marco, E. J., L. B. N. Hinkley, S. S. Hill, and S. S. Nagarajan (2011, May). Sensory Processing in Autism: A Review of Neurophysiologic Findings. *Pediatr. Res. 69*(8), 48–54.

Mayer, R. E. (2005). *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

McQuail, D. (1987). *Mass Communication Theory: An Introduction*. Sage Publications, Inc.

Milton, D. E. (2012, October). On the ontological status of autism: The 'double empathy problem'. *Disabil. Soc. 27*(6), 883–887.

Nadeem, A., A. Hilton, R. Dawes, G. Thomas, and A. Mustafa (2023). SEM-POS: Grammatically and Semantically Correct Video Captioning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2606–2616.

Neves, J. (2018). Subtitling for deaf and hard of hearing audiences: Moving forward. In *The Routledge Handbook of Audiovisual Translation*. Routledge.

Nevsky, A., F. Bircanin, E. Simperl, M. N. Cruice, and T. Neate (2025, July). To Each Their Own: Exploring Highly Personalised Audiovisual Media Accessibility Interventions with People with Aphasia. In *Proc. 2025 ACM Des. Interact. Syst. Conf.*, DIS '25, New York, NY, USA, pp. 1826–1843. Association for Computing Machinery.

Nevsky, A., T. Neate, E. Simperl, and R.-D. Vatavu (2023, August). Accessibility Research in Digital Audiovisual Media: What Has Been Achieved and What Should Be Done Next? In *Proc. 2023 ACM Int. Conf. Interact. Media Exp.*, IMX '23, New York, NY, USA, pp. 94–114. Association for Computing Machinery.

O'Nions, E., I. Petersen, J. E. J. Buckman, R. Charlton, C. Cooper, A. Corbett, F. Happé, J. Manthorpe, M. Richards, R. Saunders, C. Zanker, W. Mandy, and J. Stott (2023, June). Autism in England: Assessing underdiagnosis in a population-based cohort study of prospectively collected primary care data. *Lancet Reg. Health – Eur. 29*.

Popit, S., K. Serod, I. Locatelli, and M. Stuhec (2024, January). Prevalence of attention-deficit hyperactivity disorder (ADHD): Systematic review and meta-analysis. *Eur. Psychiatry 67*(1), e68.

Porter, B., P. Dean, N. Race, M. Lomas, and R. Ramdhany (2024, September). Self-Adaptive Systems Challenges in Delivering Object-Based Media. In *2024 IEEE Int. Conf. Auton. Comput. Self-Organ. Syst. ACSOS*, pp. 163–168.

Roman-Urrestarazu, A., R. van Kessel, C. Allison, F. E. Matthews, C. Brayne, and S. Baron-Cohen (2021, June). Association of Race/Ethnicity and Social Disadvantage With Autism Prevalence in 7 Million School Children in England. *JAMA Pediatr. 175*(6), e210054.

Romero-Fresco, P. and F. Chaume (2022). Creativity in audiovisual translation and media accessibility. *J. Spec. Transl. 38*(1), 2022.

Rong, Y., C.-J. Yang, Y. Jin, and Y. Wang (2021, May). Prevalence of attention-deficit/hyperactivity disorder in individuals with autism spectrum disorder: A meta-analysis. *Res. Autism Spectr. Disord. 83*, 101759.

Sarah Stainton (2025, November). ADHD Management Information - November 2025. https://digital.nhs.uk/data-and-information/publications/statistical/mi-adhd/november-2025.

Sardari, F., A. Mustafa, P. J. B. Jackson, and A. Hilton (2023). PAT: Position-Aware Transformer for Dense Multi-Label Action Detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2988–2997.

Sardari, F., A. Mustafa, P. J. B. Jackson, and A. Hilton (2025). CoLeaF: A Contrastive-Collaborative Learning Framework for Weakly Supervised Audio-Visual Video Parsing. In A. Leonardis, E. Ricci, S. Roth, O. Russakovsky, T. Sattler, and G. Varol (Eds.), *Comput. Vis. – ECCV 2024*, Cham, pp. 1–17. Springer Nature Switzerland.

Seltzer, M. M., M. W. Krauss, P. T. Shattuck, G. Orsmond, A. Swe, and C. Lord (2003, December). The Symptoms of Autism Spectrum Disorders in Adolescence and Adulthood. *J. Autism Dev. Disord. 33*(6), 565–581.

Shi, B., A. Tjandra, J. Hoffman, H. Wang, Y.-C. Wu, L. Gao, J. Richter, M. Le, A. Vyas, S. Chen, C. Feichtenhofer, P. Dollár, W.-N. Hsu, and A. Lee (2025, December). SAM Audio: Segment Anything in Audio. arXiv [Preprint]. arXiv:2512.18099. Available at: http://arxiv.org/abs/2512.18099 (Accessed: 2026-01-09T21:57:36Z).

Simonsen, J. and T. Robertson (2013). *Routledge International Handbook of Participatory Design*, Volume 711. Routledge New York.

Solomonidou, C., F. Garagouni-Areou, and M. Zafiropoulou (2004). Information and Communication Technologies (ICT) and Pupils with Attention Deficit Hyperactivity Disorder (ADHD) Symptoms: Do the Software and the Instruction Method Affect Their Behavior? *J. Educ. Multimed. Hypermedia 13*(2), 109–128.

Song, H.-L., J.-D. Kong, Q.-L. Wang, and C.-X. Zhou (2025, April). Gender identity and its impact on attention deficit hyperactivity disorder manifestations: Challenges in diagnosis and assessment. *World J. Psychiatry 15*(4), 101952.

Sowerby, P., S. Seal, and G. Tripp (2011, August). Working Memory Deficits in ADHD: The Contribution of Age, Learning/Language Difficulties, and Task Parameters. *J. Atten. Disord. 15*(6), 461–472.

Spiel, K., E. Hornecker, R. M. Williams, and J. Good (2022, April). ADHD and Technology Research – Investigated by Neurodivergent Readers. In *Proc. 2022 CHI Conf. Hum. Factors Comput. Syst.*, CHI '22, New York, NY, USA, pp. 1–21. Association for Computing Machinery.

Stevens, T., L. Peng, and L. Barnard-Brak (2016, November). The comorbidity of ADHD in children diagnosed with autism spectrum disorder. *Res. Autism Spectr. Disord. 31*, 11–18.

Straub, L., B. T. Bateman, S. Hernandez-Diaz, C. York, B. Lester, K. L. Wisner, C. J. McDougle, P. B. Pennell, K. J. Gray, Y. Zhu, E. A. Suarez, H. Mogun, and K. F. Huybrechts (2022, March). Neurodevelopmental Disorders Among Publicly or Privately Insured Children in the United States. *JAMA Psychiatry 79*(3), 232–242.

Team, S. D., X. Chen, F.-J. Chu, P. Gleize, K. J. Liang, A. Sax, H. Tang, W. Wang, M. Guo, T. Hardin, X. Li, A. Lin, J. Liu, Z. Ma, A. Sagar, B. Song, X. Wang, J. Yang, B. Zhang, P. Dollár, G. Gkioxari, M. Feiszli, and J. Malik (2025, November). SAM 3D: 3Dfy Anything in Images. arXiv [Preprint]. arXiv:2511.16624. Available at: http://arxiv.org/abs/2511.16624 (Accessed: 2026-01-09T21:56:49Z).

Ward, L., M. Paradis, B. Shirley, L. Russon, R. Moore, and R. Davies (2019, October). Casualty Accessible and Enhanced (A&E) Audio: Trialling Object-Based Accessible TV Audio. In *147th Audio Eng. Soc. Conv.*, New York, NY, USA.

Ward, L., B. Shirley, and J. Francombe (2018, October). Accessible Object-Based Audio Using Hierarchical Narrative Importance Metadata. In *145th Audio Eng. Soc. Conv.*, New York, NY, USA.

Zabalbeascoa, P. (2008, September). The Nature of the Audiovisual Text and Its Parameters. In J. Díaz-Cintas (Ed.), *The Didactics of Audiovisual Translation*, pp. 21–37. John Benjamins Publishing Company.

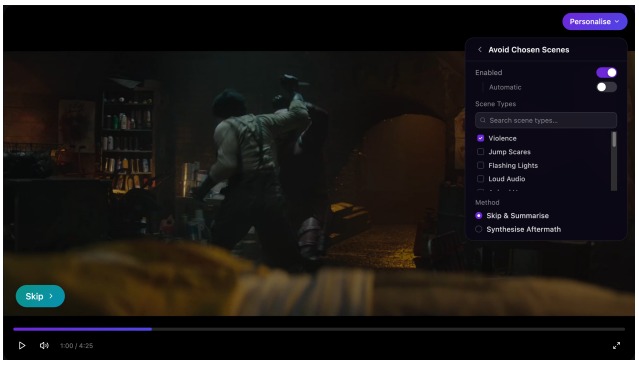
Zárate, S. (2021). *Captioning and Subtitling for d/Deaf and Hard of Hearing Audiences*. UCL Press.

Zhang, Y., A. Caliskan, M. Xu, A. Hilton, and J.-Y. Guillemaut (2025). MVL-Net: Pairwise Learning for Multi-View Multiple People Labelling. *IEEE Trans. Multimed. 27*, 4736–4751.

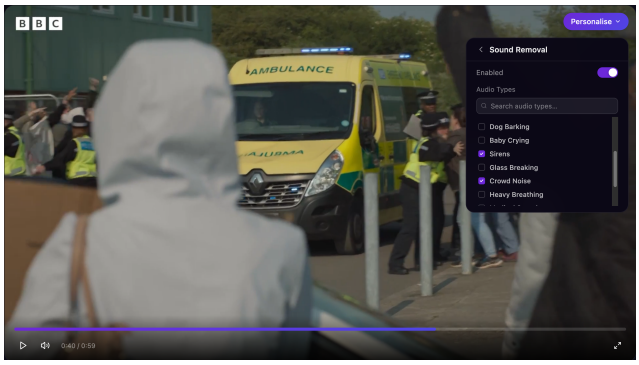
Zhu, H. H., R. Chen, and Y. Zhao (2025, October). FocusView: Understanding and Customizing Informational Video Watching Experiences for Viewers with ADHD. In *Proc. 27th Int. ACM SIGACCESS Conf. Comput. Access.*, ASSETS '25, New York, NY, USA, pp. 1–18. Association for Computing Machinery.

A. INTERACTIVE INTERVENTION MOCKUPS

Figure 4 presents the interactive intervention mockups.



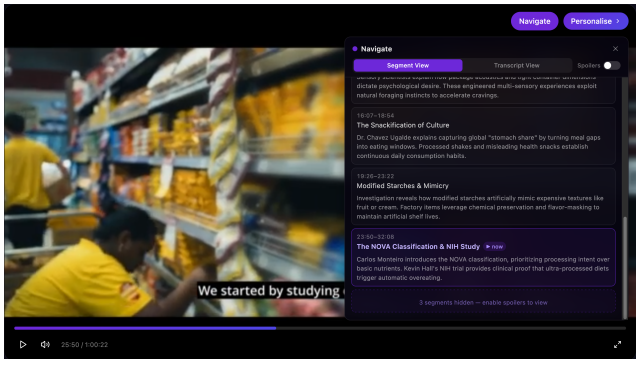
(a) *Scene Classification for Empowered Watching Choice*



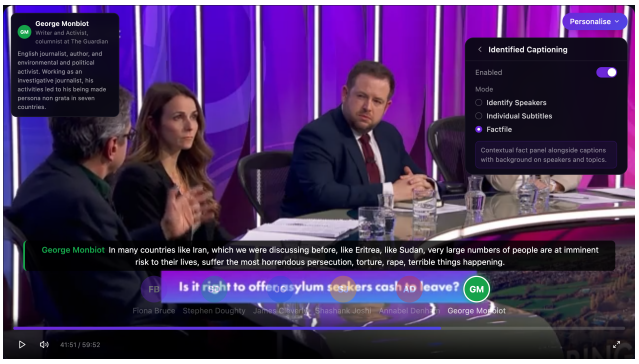
(b) *Sound Replacement or Removal to Reduce Discomfort*



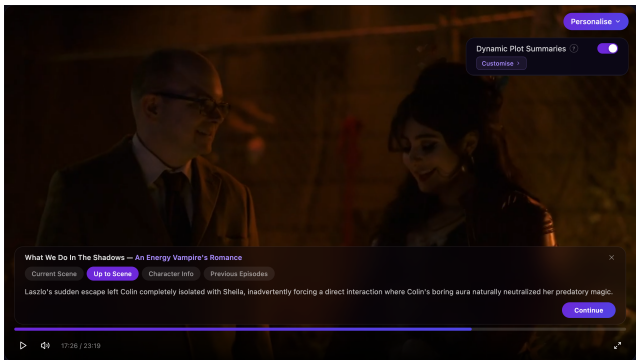
(c) *Sound Remixing to Reduce Distraction*



(d) *Content Segmentation for Efficient Navigation*



(e) *Identified Speaker Captioning to Improve Comprehension of Complex Dialogue*



(f) *Dynamic Plot Summaries to Enrich Narrative Understanding*

Figure 4: *Screenshots of the interactive intervention mockups created in Figma Make, corresponding to the six functional intervention clusters detailed in Section 5.2 and evaluated in Section 6.*