

Embodied Relaxation among Adolescent ASMR Viewers on Bilibili: An Exploratory Mixed-Methods Study

Ke Qu

*Guangdong Experimental High School (Liwan Campus), Guangzhou, China
qukelia0811@163.com*

Abstract. This study examines embodied relaxation among adolescent viewers of Autonomous Sensory Meridian Response (ASMR) content on Bilibili, focusing on how audio-visual cues are interpreted by the audience and on the para-social interaction mechanisms that arise from them. An exploratory mixed-method design was adopted, combining a questionnaire with semi-structured interviews and approaching the material from the audience-interpretation perspective. Seventeen valid questionnaire responses were obtained through convenience and snowball sampling among viewers who were mainly aged 12 to 18, and a purposive subsample of respondents who volunteered was then interviewed. Questionnaire results indicate that low-intensity soft lighting, warm color palettes, minimalist solid-color backgrounds, close-up hand movements and cute plush props are widely accepted, whereas intense close-up shots that create direct eye contact carry individual acceptance thresholds and were reported as uncomfortable by a minority of respondents. Tapping sounds were treated as the core relaxation element by nearly half of the respondents, while a further group weighted audio and visual cues roughly equally, and soft human voices and ambient sounds were valued mainly for anxiety relief and sleep aid. Interview accounts suggest that ASMR is used as an active cognitive reappraisal strategy, and that role-play content builds one-sided intimacy through close-range audio-visual cues. Given the small sample, these findings are exploratory rather than confirmatory. The study applies emotion-regulation and parasocial-interaction theory to sensory short-video content and offers references for healthy ASMR content creation.

Keywords: ASMR, emotion-regulation theory, parasocial-interaction theory, embodied relaxation, relieve anxiety

1. Introduction and literature review

Autonomous Sensory Meridian Response (ASMR) can be defined as the subjective experience with low level of euphoria. It also has the characteristics of combining the feelings of excitement and static-like tingling sensation on the skin. Although ASMR can occasionally be induced by intentional attention control, it is most reliably and commonly triggered by specific auditory and visual stimuli [1, 2]. Typical triggering elements include soft whispering, tapping, rubbing sounds, personal attention, and simulated touching. ASMR is primarily disseminated through video platforms, fostering a culture of digital embodiment. Chinese video platform Bilibili has become a

central community where young users consume and create ASMR content. The content style on Bilibili integrates local aesthetics such as "healing", "slow living", and "immersive experience" [1-3].

The core analytical framework of Gross's emotion regulation theory [4] draws difference between two primary tactics, antecedent-focused regulation (cognitive reappraisal) and response-focused regulation (expressive suppression). After the process of cognitive reappraisal alters the personal thinking about a situation, an emotional response will be fully generated. It can reduce the adverse emotional impact [5]. Expressive suppression refers to the restraint of ongoing emotion-displaying actions once an emotion has already been elicited [4].

ASMR culture in videos can provide the sense of sensory companionship without pressure [2], and the initial application of ASMR culture suggests that ASMR is regarded as a proactive cognitive reappraisal strategy. Individuals with a stronger cognitive reappraisal tendency can have good experience [6].

Although most existing studies focus on the physiological and psychological effects of ASMR, such as changes in heart rate, skin conductance and perceived relaxation, existing studies seldom interpreted how online users interpret auditory and visual elements and how these elements provoke the sense of embodied relaxation [7]. That is the first research gap which will be addressed in this study. Existing research covers a broad age spectrum, yet few studies specifically study how teenagers aged 12–18 interpret ASMR elements. That is the second research gap, which will be addressed in this study. Additionally, few scholars pay attention to the unique mechanism of ASMR videos, in which performers cannot receive real-time feedback from viewers and which can be seen as the one-way communication depending on the information output of performers. Para-social interaction in this case can be interpreted as the one-sided emotional bond of intimacy perceived solely by the viewers [8]. Lastly, there are ambiguous boundaries between ASMR and other healing-related content. Bilibili hosts abundant hybrid videos integrating immersive makeup and slow-living vlogs. The conclusion of this study will address this gap based on empirical findings.

This study addresses three research questions (RQs).

RQ1 asks which visual elements in Bilibili ASMR videos young users find most relaxing or immersive.

RQ2 asks how young Chinese users interpret common ASMR sound elements, such as whispering, tapping, soft sounds and personal attention.

RQ3 asks how young viewers relate specific parasocial cues, such as a performer's smiling, sustained eye contact, use of "you" and gentle scolding, to their sense of being cared for and emotional comfort.

2. Method

2.1. Research design

This study adopted a mixed method research design. Firstly, the quantitative questionnaire can be used to collect statistical data regarding Bilibili young users' overall preferences for audiovisual elements, viewing frequency and perceived relaxation. It can help to address the RQ1. Next, using qualitative semi-structured interviews invited respondents being selected among the questionnaire participants. It can be used for exploring their subjective interpretation and emotional mechanisms behind para-social interaction. That can be used for answering RQ2 and RQ3 uncovering the underlying psychological and cultural insights that quantitative questionnaires cannot explain.

2.2. Participants and sampling

In terms of target participants, they are selected among the platform users. The inclusive criteria of participants can be defined as long-term viewers of ASMR content on Bilibili. The age difference has also been considered. Participants mainly include teenagers aged 12–18. Participants are also required to have the experience of fully watching more than 3 ASMR or slow-lifestyle hybrid videos published on Bilibili.

The sampling method for questionnaire is convenience sampling on WeChat group and snowball sampling. Participants with close relationship helped to introduce more ASMR viewers to do questionnaire. The valid sample consists of 17 respondents, and all proportions are calculated on this base. Participants are stratified by viewing frequency into occasional viewers ($n = 10$, 58.8%), users watching 1–3 times weekly ($n = 3$, 17.6%), daily viewers ($n = 2$, 11.8%), and users who rarely or never watch relevant content ($n = 2$, 11.8%).

Purposive sampling has been used for selecting interviewees. The interviewees can be selected from participants of questionnaire who stated that they are willing to participate in in-depth interviews. Three screening criteria were applied to ensure the diversity of interviewees. The first criterion covers two viewing behaviour groups, namely high-frequency heavy users and casual occasional users. The second criterion covers divergent experiential types, namely participants who feel comfortable and participants who feel embarrassed with close-up direct camera shots. The third criterion covers distinct content preferences, namely preference for tapping sounds, for whispered vocal audio, and for immersive mixed-format content.

2.3. Questionnaire design and procedure

Step 1 is the design of the questionnaire. Firstly, the basic information module included age group, gender, and frequency of watching Bilibili ASMR content (daily, 1–3 times per week, 1–3 times per month, rarely watch, never watch). Secondly, the visual element perception module used a 5-point Likert scale ranging from no relaxation at all to extreme relaxation, and its measured variables covered solid minimalist background, close-up handheld shots, warm color palette, low-intensity soft lighting, cute plush props, and direct camera framing. Thirdly, the auditory sound perception module combined Likert scale and single-choice questions, and its measured variables covered whispered human voice, tapping sound effects, and gentle ambient sound; this section quantifies the respective weight of visual and auditory stimuli on viewers' perceived relaxation. Fourthly, the audiovisual synchronization and content category module assessed constructs covering intimacy generated by paired close-range audio capture and close-up footage, and audience perception of content boundaries between pure ASMR and immersive vlog hybrid formats. Fifthly, an open-ended closing question asked about the willingness to participate in follow-up interviews.

Step 2 is the online distribution and data collection process. Before the formal distribution of questionnaire, reliability and validity tests were conducted for this self-developed ASMR audiovisual perception questionnaire to secure robust measurement quality. The test of validity has been developed through inviting two experts specializing in psychology and digital media to review all question wording, revise ambiguous items, in order to improve construct validity. Cronbach's value achieved $\alpha > 0.7$, indicating acceptable reliability.

Data collection was mainly implemented via online questionnaires distributed through Sojump. Invalid submissions (entries with excessively short completion time, uniform answers across all items) were excluded from analysis.

2.4. Interview design and procedure

The interview opens with introductory warm-up questions to clarify interviewees' viewing frequency, viewing scenarios, viewing motivations and content preferences, so as to establish a holistic understanding of interviewees' basic usage behaviors. These warm-up questions also record how long interviewees have had access to ASMR content and which platform channels they habitually use.

The core audio section raises questions around typical ASMR auditory stimuli including whispering, tapping and friction sounds, as well as the exclusive sense of intimacy brought by close-range sound pickup. It explores users' subjective perceptions of different auditory stimuli, addressing the research questions about audio element interpretation in RQ2 and RQ3.

The visual section focuses on visual components such as background lighting and shadow, close-up shots, direct camera angles and cute plush props. It asks about how elements affect users' immersion and relaxation experience, supplements questionnaire data, and identifies the visual symbols preferred by audiences.

Throughout the interview, questions are arranged regarding interviewees' perceptions of para-social interaction cues such as creators' smiles and eye contact. It explores how such cues generate audiences' feeling of being cared for, and answers the research question on para-social interaction proposed in RQ3.

In terms of the interview implementation process, online audio or video interviews are adopted, with each session lasting 30 to 60 minutes. All interviews are recorded, and full transcripts are produced after each interview.

2.5. Data analysis

Quantitative questionnaire data were processed with Excel and SPSS, adopting descriptive data analysis to calculate the proportion of audience acceptance for each audiovisual element and the distribution of viewing frequency. The analytical outputs extract audiences' widely preferred visual elements, including low-luminance soft lighting, warm color palettes, solid-color backgrounds, close-up hand shots and cute plush props, together with core auditory demands for gentle human voices and ambient sounds serving sleep aid and stress relief. Interview findings are reported at the level of recurrent themes rather than as counts of interviewees.

2.6. Ethical considerations

Interviewees were drawn from questionnaire respondents who had stated their willingness to take part in a follow-up interview. Before interviews, researchers inform participants about the research purpose and about anonymization, and obtain informed consent for audio recording. Interviewees' names are not disclosed, and all identifiable personal information will be removed from transcripts. The study collected no physiological measures.

3. Questionnaire results

In terms of the viewing frequency of ASMR videos, some of the viewers ($n = 3$, 17.6%) watch ASMR 3 times per week, while daily viewers account for $n = 2$ (11.8%). Heavy loyal users only constitute $n = 5$ (29.4%) of the sample. One respondent rarely watches such content, and there is only one who has never watched ASMR videos at all, each accounting for 5.9%. Overall, more than

half of the respondents only browse ASMR content casually in their free time, while regular high-frequency viewers make up a small proportion.

As the results related to the impact of visual elements on relaxation and immersion reveal, more than 70% of respondents stated that minimalist solid-color backgrounds can bring them a sense of relaxation. There are $n = 3$ (17.6%) who felt a strong relaxing effect and $n = 9$ (52.9%) who felt a slight relaxing effect, accounting together for $n = 12$ (70.6%). Only a small group of respondents have no feelings of relaxation or even negative reactions to such backgrounds. It shows that clean, minimalist visuals well match these respondents' demand for relaxation.

A further $n = 9$ (52.9%) of users believed close-up shots could moderately improve their sense of immersion, while $n = 5$ (29.4%) reported no obvious impact. None of the users stated that close-ups brought a dramatic improvement to immersion. This indicates that close-up footage serves as a bonus element rather than a decisive factor for immersive experience.

A total of $n = 12$ (70.6%) of users accepted the relaxing effect of warm color palettes. Among them, $n = 2$ (11.8%) consider the effect excellent and $n = 10$ (58.8%) think it fairly good. No users considered the effect poor, which suggests warm hues are these respondents' preferred visual color scheme.

Hand movements were recognized as the core immersive visual element by $n = 10$ (58.8%) of users, comprising $n = 2$ (11.8%) who fully agreed and $n = 8$ (47.1%) who partially agreed. A further $n = 5$ (29.4%) of the sample remained uncertain. Views of hand manipulation act as a core visual symbol for building immersion.

A proportion of $n = 13$ (76.5%) of users agreed that pictures with dim-light visuals bring physical and mental soothing feelings. There are $n = 2$ (11.8%) who strongly agreed and $n = 11$ (64.7%) who agreed, with zero negative feedback. Soft, low-lit ambiance is highly accepted by these respondents.

A further $n = 11$ (64.7%) of participants thought cute soft props match the relaxing atmosphere and enhance the experience of immersion, while only $n = 1$ (5.9%) found props distracting. Gentle lovely props fit the non-stressful and relaxed meaning of ASMR.

A proportion of $n = 13$ (76.5%) of respondents gained immersive feelings from the shot with the strong feeling of eye contact, combining $n = 5$ (29.4%) who felt strongly immersed and $n = 8$ (47.1%) who felt slightly immersed. The remaining $n = 4$ (23.5%) felt awkward or disgusted. Different audiences have diverse feelings of intimate companion.

To sum up, within all visual dimensions, low-brightness soft lighting, warm color palettes, minimalist solid-color backdrops, soft cute props, and close-up hand movements are widely favored visual elements among these respondents, which these respondents reported as improving relaxation and immersion. However, intense close-up eye contact shots come with a certain threshold for audience who can accept and may trigger discomfort among a portion of users.

In terms of auditory sound's emotional experience, $n = 8$ (47.1%) of users suggest tapping sound effects as the core element for relaxation, while visuals only serve as auxiliary tools to build immersion. Another $n = 7$ (41.2%) pay more attention to the footage of hand tapping movements, showing that audio and visual elements carry nearly equal weight for this group and tapping-style ASMR content features a strong audio-visual binding effect. A proportion of $n = 12$ (70.6%) of users believe soft human voices and ambient sounds have emotional soothing value, combining $n = 8$ (47.1%) who reported anxiety relief and sleep aid with $n = 4$ (23.5%) who reported emotional comfort, which is the most frequently reported demand driving these respondents to listen to ASMR videos.

For audio-visual synergy perception, the combination of close-range audio recording and close-up footage delivers an intimate healing experience for nearly half of the respondents, yet a small

subset of them still reports discomfort with this style.

Overall, the creation of ASMR video content should put the soft, therapeutic audio design into the first place and be paired with low-saturation warm lighting and clean, minimalist visuals. Creators are advised to avoid intense close-up eye contact shots to catch the relaxation needs of most respondents in this sample.

4. Interview results

4.1. Motivations for overall ASMR-video consumption and interviewees' interpretation of ASMR

As interviewees demonstrated in interview, they firstly encountered ASMR on Bilibili during 2017–2018. With the development of new media, algorithm-driven platform recommendations have become the primary access channel.

According to the results of online questionnaire, only $n = 5$ (29.4%) of questionnaire respondents are heavy users with daily or weekly high-frequency consumption. Among the interviewees, those falling into this heavy-user category report daily listening habits and serve as a representative case for high-frequency users.

In terms of the usage scenario, interviewees habitually wear headphones and listen to the sound in ASMR videos before bedtime, forming a long-term lifestyle routine. This differs from light-user respondents in the questionnaire who "browse the ASMR videos casually in fragmented spare time".

According to interviewees' interpretations, ASMR is fundamentally defined as a sleep-aid medium; functional benefits take priority, while entertainment value is secondary. This aligns with Gross's emotion regulation theory, which distinguishes antecedent-focused regulation (cognitive reappraisal) from response-focused regulation (expressive suppression). ASMR functions as an emotion-regulation tool that audience take initiative to use it.

4.2. Interviewees' preference for auditory elements

The analysis of interview data related to auditory-stimulation preferences can be used for addressing RQ2 and explaining interviewees' interpretation of various ASMR audio elements. The most attractive auditory elements for interviewees are ethereal wooden tapping sounds, sound for tapping on plastic casings of nail-art and cosmetic products, and soft mouth-generated human sounds. These results can be consistent with the results of online questionnaire, showing that $n = 8$ (47.1%) of survey participants identify tapping sounds as core relaxation triggers, and $n = 12$ (70.6%) acknowledge the soothing effect of soft human vocalizations. Interviewees' interpretations are broadly in line with the preferences reported by the survey respondents.

Moreover, interviewees identified the evaluation criteria for high-quality audio elements, namely low background noise and moderate human-voice volume without excessively loud nor faint volume. They all agree that the sound elements with the feeling of comfort should be the first choice.

In terms of the integration of visual and auditory cues, sound constitutes the primary element, while hand-movement footage serves merely as supplementary elements. This matches the questionnaire finding that nearly half of the respondents regard tapping sounds as the core and visuals as secondary.

4.3. Perception of immersive experience of visual elements

In terms of high-value visual guiding elements recognized by the interviewees, light-guided shots and shots featuring hand-gaze guidance with props like pens and nail-art tools can enhance the experience of immersion. Low-intensity soft light, warm colour tones, minimalist solid-colour backgrounds, close-ups of hands, and cute props are visually preferred by the respondents in general. The interview case supplements the significance of the gesture for guidance by ASMR creators, which can be regarded as the specific type of visual cue.

In addition, respondents favour close-up shots that bring role-playing performers closer to the audience. Nevertheless, such shots may be rejected by some users. The respondents acknowledge that the combination of close-range sound recording and hand close-ups boosts a sense of comfort and intimacy. This aligns with the questionnaire result that nearly half of the respondents favour the close-range audio and visual elements.

4.4. Parasocial interaction between online users and the creator of ASMR video

Preference is given to role-playing content with human vocal tracks, rather than purely ambient videos consisting only of hand close-ups. This scenario can be defined as a kind of parasocial interaction in which, while watching, the respondent voluntarily identifies with the recipient role and passively receives companionship and comfort from creators. To be specific, gentle vocal delivery from creators and caregiving gestures captured in close proximity evoke the emotional demand of respondents that desire for being cared and accompanied.

5. Discussion and conclusion

5.1. Summary of findings

In terms of users' audio-visual preference, the young Bilibili users surveyed here prefer dim, soft-light visuals with warm-toned minimal styles. Tapping sounds and soft human voices serve as key stimuli that bring the feeling of relaxation. Content creation should follow the principle that audio stimuli in video clips should prioritize the users' preference, and visual elements can support the demonstration of relaxed and comfortable feelings. Whether audiences can accept shots filled with direct close-up eye-contact depends on different individuals' preferences.

While discussing the mood regulation and ASMR value, the findings indicate that ASMR is used as a cognitive reappraisal tool. Respondents report that ASMR videos help them become relaxed when they feel anxious before exams and ease their emotional distress. The heavy users interviewed here develop long-term bedtime listening habits, and ASMR acts as an important sensory companion in their daily life.

Furthermore, according to the interview accounts, role-play content in ASMR videos builds one-way social intimacy through close-range audio-visual cues. Users place themselves in the recipient role and gain mental comfort without real-world social pressure.

5.2. Theoretical contributions

Based on Gross's emotion-regulation theory, this study suggests that ASMR audio-visual sensory stimulation functions as a spontaneous everyday coping strategy. In addition, this study applies the theory of para-social interaction in the digital age to interpret how close-range audio-visual cues within short-form videos may produce a virtual sense of being cared for.

5.3. Limitations

Sampling bias of this study can be described as the application of convenience sampling in questionnaire, resulting in the limited representativeness of sample. The valid questionnaire sample is small, with 17 respondents, so the proportions reported above describe this group only and are not generalized to adolescent ASMR audiences as a whole. The interview subsample is likewise small and was drawn from questionnaire respondents who volunteered, and the number of interviewees is not reported in this study, so interview findings are reported as recurrent themes rather than as counts. The sample was moreover composed mainly of viewers aged 12 to 18, which further limits how far the findings can be extended to other age groups. Response distributions are reported only for the categories summarised in the original analysis, and they do not always sum to the full sample. The response anchors appearing in the results do not in every case correspond to the single Likert scale described in the method, and the full item wording of the questionnaire is not reported. Reliability was reported only as an overall coefficient above the conventional threshold, without values for individual subscales. The questionnaire collected age group and gender, but neither variable is reported in the results, so no subgroup comparison is available. No approval from an institutional ethics review board is reported for this study, nor is consent from the guardians of participants under 18, although the reported procedure includes informed consent for audio recording and the removal of identifiable personal information from transcripts. In addition, confounding variables such as listening duration, headphone devices and individual anxiety levels were not controlled. Moreover, no physiological measures like heart rate were collected. Results only rely on self-reports, so relaxation effects cannot be objectively verified.

5.4. Future research suggestions

There are several future research directions. First, interview samples can be expanded to balance teenagers aged 12–18 and young adults, and to systematically compare their differences in audio-visual perception. Second, experimental methods with physiological monitoring like heart rate can objectively measure ASMR's relaxing effects, instead of relying merely on subjective questionnaires. Third, participants can be divided into heavy-use, light-use and occasional-use groups, so that mood-regulation outcomes across different usage frequencies can be compared. Fourth, the creator perspective in ASMR production can be explored, and preference differences between creators and viewers can be compared.

References

- [1] Ahuja, N. K. (2013). "It feels good to be measured": Clinical role-play, Walker Percy, and the tingles. *Perspectives in Biology and Medicine*, 56(3), 442–451. <https://doi.org/10.1353/pbm.2013.0022>
- [2] Trenholm-Jensen, E. A., Burns, L., Trenholm, J. E., & Hand, C. J. (2022). Beyond tingles: An exploratory qualitative study of the Autonomous Sensory Meridian Response (ASMR). *PLOS ONE*, 17(12), e0277962. <https://doi.org/10.1371/journal.pone.0277962>
- [3] Starr, R. L., Wang, T., & Go, C. (2020). Sexuality vs. sensuality: The multimodal construction of affective stance in Chinese ASMR performances. *Journal of Sociolinguistics*, 24(4), 492–513. <https://doi.org/10.1111/josl.12410>
- [4] Gross, J. J. (1999). Emotion regulation: Past, present, future. *Cognition and Emotion*, 13(5), 551–573. <https://doi.org/10.1080/026999399379186>
- [5] Li, J. (2024). "I feel soothed and cured": The effects of ASMR and how HCI research can benefit from them. *Interactions*, 31(5), 12–14. <https://doi.org/10.1145/3679052>
- [6] Fanning, E. L. (2026). Harnessing ASMR: A guide to reducing stress through sensory experiences in corporate wellness programs [Doctoral dissertation, University of Arizona Global Campus]. *ProQuest Dissertations and*

Theses (No. 32737255).

- [7] Du, J., You, L., & Zhang, J. J. (2026). Exploring the therapeutic potential of VR-based ASMR animation: A comparative study on relaxation and sleep aid. *Computer Animation and Virtual Worlds*, 37(2), e70101. <https://doi.org/10.1002/cav.70101>
- [8] Stock, W. G., Fietkiewicz, K. J., Scheibe, K., & Zimmer, F. (2022). Cyber social interactions: Information behavior in between social and parasocial interactions. *Journal of Information Science Theory and Practice*, 10(3), 15–23. <https://doi.org/10.1633/JISTaP.2022.10.3.2>