

87 Hairstyles Are Not Enough: How Players Experience Character Creation in *Baldur's Gate 3*

MARCO WEISS*, Saarland University, Saarland Informatics Campus, Germany

PASCAL LESSEL*, German Research Center for Artificial Intelligence (DFKI), Saarland Informatics Campus, Germany



Fig. 1. Composite of three screenshots from the character editor of *Baldur's Gate 3* by Larian Studios [G1], annotated by the authors to illustrate the interface concepts of customization categories (red boxes) and customization options (blue boxes) used throughout this paper.

In digital roleplaying games character editors are commonly present. Here, not only visual but often also gameplay-related decisions are made. In this paper, we focus on the fantasy roleplaying game *Baldur's Gate 3*. Through an online questionnaire ($N = 171$), we investigated how the character editor is perceived, which customization categories players find important and how they assess the number of options offered within categories. We derived design suggestions and our results show that the editor is perceived well overall, but can improve, mainly by offering more categories and more selection options in the existing categories, even in those, that already offer many options. In contrast, categories that already had many options (100+) did not overwhelm participants, indicating that it is not a general problem to offer many options. Complementary to this, we found relationships between individual differences (age, gender, personality and engagement with the game) and the category importance perception, opening up questions around tailoring character editors.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **User studies**.

*Both authors contributed equally to this research.

Authors' Contact Information: [Marco Weiss, mawe00007@stud.uni-saarland.de](mailto:mawe00007@stud.uni-saarland.de), Saarland University, Saarland Informatics Campus, Saarbrücken, Saarland, Germany; [Pascal Lessel, pascal.lessel@dfki.de](mailto:pascal.lessel@dfki.de), German Research Center for Artificial Intelligence (DFKI), Saarland Informatics Campus, Saarbrücken, Saarland, Germany.



This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2026 Copyright held by the owner/author(s).

ACM 2573-0142/2026/11-ARTGAMES005

<https://doi.org/10.1145/3831278>

Additional Key Words and Phrases: Avatar Creation, Avatar Design, Character Creation, Character Editor, Roleplaying Games, RPG, Customization, Individualization, Individual Differences

ACM Reference Format:

Marco Weiss and Pascal Lessel. 2026. 87 Hairstyles Are Not Enough: How Players Experience Character Creation in *Baldur's Gate 3*. *Proc. ACM Hum.-Comput. Interact.* 10, 7, Article GAMES005 (November 2026), 35 pages. <https://doi.org/10.1145/3831278>

1 Introduction

Character¹ editors are common in many interactive contexts, such as video games, virtual worlds, social platforms, and online communities where users can create and customize characters [1, 6, 11, 12, 40]. They allow users to create digital representations in the form of characters and act as an interface between the user and the virtual environment. Considering choices, editors consist of customization categories (e.g., *Hairstyle*, *Scarring*; see Figure 1, red boxes) and selection options within these (e.g., which kind of scar, see Figure 1, blue boxes). Depending on the usage context, editors offer the possibility to modify aesthetic aspects, allow users to influence properties beyond appearance, or both. For example, in video role-playing games² (RPGs), they often allow both, and some decisions in the editor impact the player options in the game.

Research in the context of games shows that character customization can influence identification with the character or the game's immersion [5, 6] and work done in social and virtual environments indicates that visual self-representation within the environment impacts perception, behavior, and social interaction [12]. Character editors are therefore more than configuration tools as character construction and similarity can also affect presence, perceived control, and performance [35]. To our knowledge, there is a lack of larger-scale user-centered studies on character editors that investigate which features and customization categories are important for users and how many options they require. This is surprising, given that character editors come with an inherent, yet to our knowledge not understood, tension: On the one hand, they need to offer (supposedly) many options and fine-grained control to enable every user to create desired characters, and depending on the usage context, these decisions need to be made up-front and would have long-lasting effects in the game, as these are not necessarily (fully) adjustable in later stages of the experience³. On the other hand, offering too many options might lead to choice overload [20] and trade-offs between usability and expressiveness [40], raises the questions how users are able to navigate/explore their available options. Therefore, character editors as a customization tool are an application domain in HCI to understand how users tolerate and even demand complexity when it serves expressive goals.

Previous works have considered editors, but, often without involving the players' perspective (e.g., [29, 33]), for example, by doing an author assessment only. Studies involving users and analyzing their preferences for the available choices within complex editors, to our knowledge, are scarce. Ducheneaut et al. [5] addressed related questions in three virtual worlds in 2009, but these editors offered fewer, less fine-grained customization categories than many contemporary editors do. More recent user-centered studies, such as McArthur [28], have examined character customization practices and creation context but have not systematically assessed the importance of customization categories and the sufficiency of options across a broad set of editor categories. To address this gap, we investigate the video game context, as editors here typically require complex decisions and are

¹In this work, the term *character* is used synonymously with *avatar* as the examined video game also uses the term *character*.

²A game genre in which players assume and act out the roles of characters in a fictional setting, cf. <https://www.universityxp.com/blog/2025/3/24/what-are-role-playing-games>, last accessed: 07.07.2026

³Although some games allow adjustments, even popular ones might disallow or restrict these: for example, *Cyberpunk 2077* [G3] only allows partial adaptations during gameplay. Thus, if players are unhappy with particular initial choices, they need to restart the game.

often offering aesthetic and game-related customization categories. To obtain opinions from many players, we decided to conduct an online survey (which was mainly answered by participants from Germany and the US). Here, we focused on one game to ensure that all participants have the same context and can assess the same known categories, instead of having many participants that all have different games with different editors in mind. Genre-wise, we assessed the RPG genre as well suited, as usually the player's own character is put at the center of the experience and prior work on identity-related self-expression in games often highlights RPGs as a relevant context [25]. As a representative, modern and high-production-value game with a large player base (to reach enough potential participants) we selected *Baldur's Gate 3* (BG3) [G1], which mechanically is based on the popular fantasy tabletop RPG Dungeons & Dragons (D&D): At the time of our study (July 2025), BG3 had an average daily player base of over 80,000⁴ and was praised by players as well as critics⁵. Scientifically, it has also been considered for other topics, such as progression balancing [34], queer representation in games [14], and music analysis [46]. BG is also relevant for representation as it avoids strict gender-locking, to not constrain non-binary players [42]. While the game also allows players to pick a pre-defined character with few customization options ("origin characters") or fully randomized ones, we are focusing on the full ("custom") character creation. Also, we do not cover how a character evolves during the game. We address these research questions:

RQ 1 How do players perceive the BG3 character editor and do they identify shortcomings?

RQ 2 How do players evaluate the importance of individual customization categories, and is the number of options within these categories sufficient for them?

RQ 3 How are individual differences related to the perception of the editor its categories?

Learning about shortcomings of the BG3 character editor in **RQ 1** helps to situate our results. **RQ 2** revisits work from 2009 on category importance and option sufficiency [5] in a contemporary, more fine-grained editor to understand the aforementioned tension. As work showed that personality and gender impact resulting characters [5, 6, 12], with **RQ 3** we investigate whether this extends also to perception of customization categories. Our major findings are:

- (1) Our data shows that while the BG3 character editor is perceived well overall, it needs more customization categories and options to satisfy every player. We also found that players are not overwhelmed by a large number of options to choose from, suggesting that options could be increased significantly without harm (given a proper user interface design). This also seems necessary, as even categories that already offer many selection options (e.g., 87 *Hairstyles* or 45 *Tattoos*), were assessed to need more options by the majority of the sample.
- (2) We show that the majority of the categories is important (and we provide reasons why) for the majority of our sample and that no category exists that does not matter for any player. Categories having an impact on the game (i.e., mechanically; available choices in dialog; or non-player character (NPC) responses), were rated as important, suggesting that it might be worthwhile in the future to give certain—today only purely aesthetic—categories an impact on the gameplay as well (e.g., NPC dialog changes acknowledging certain hair colors). However, our data also shows that *Hairstyle* is the most important purely cosmetic category, with ratings similar to the most important game-related categories, indicating that it is not mandatory to give every category a direct game impact. We complement the work of Ducheneaut et al. [5] from 2009, showing that the importance of certain aesthetic elements is still similar today.
- (3) We found that age, gender, personality and engagement time with the game are correlated to the perception of certain customization categories.

⁴<https://steamdb.info/app/1086940/charts>, last accessed: 07.07.2026

⁵For example, see <https://www.metacritic.com/game/baldurs-gate-3>, last accessed: 07.07.2026

Overall, through the results and design suggestions this paper contributes guidance for the development of character editors (e.g., when development resources are limited, which categories should be focused on) and that effective character editors should not limit choices for player's expressiveness, either through a too-small number of options or by implementing rules to limit the selection in specific categories based on previous choices. It also shows that the assumed tension is not present, or that choice overload in this context begins at much higher numbers of options. Finally, based on the found individual perception differences, our data suggests that it is reasonable to start researching individualized editors to move away from today's "one size fits all" approach.

2 Related Work

Two topics are particularly relevant: character editors as customization systems and user character.

2.1 Character Editors

Related work explores character editors through design and user-centered approaches

2.1.1 Interface- and Affordance-focused Analyses. McArthur and Jenson [29] systematically analyzed the interface of character editors using a qualitative, design-oriented approach to study character creation interfaces in video games (*Mii Maker* [G8], *World of Warcraft* (WoW) [G14], *Saints Row 2* [G10], *Jam City Rollergirls* [G6]). They developed a tool called the *Avatar Affordance Framework* to structure this analysis. The authors conducted the analysis from an expert perspective and did not include empirical validation with users. They show that categories, interface widgets, hierarchies, and default values shape available decision options and argue that the design of character editors is not neutral: it facilitates/hinders forms of expression through restricted categories or preset default values. For example, the binary representation of identity (female/male) could prevent the creation of the desired character or prevent non-binary individuals from expressing themselves. Another limitation is a limited range of body types through pre-made presets. More recent work focused on gender expression and body representation. Løfgren Elvig and Mekler [8] analyzed 19 character editors, including *BG3*'s character editor, to investigate how they afford non-binary and gender expression. They used the walkthrough method [23] for data collection and combined feature analysis [17] to examine both explicit customization features and the cultural meanings embedded in the interface. They demonstrated that customization categories, such as pronouns, body type, facial hair, and genital options, can enable, constrain, or complicate gender expression. They also showed that even gender-neutral options may be implemented in ways that make them less accessible, for example, in *BG3*, gender-neutral pronouns are only accessible after navigating through the *Edit Appearance* tab. While their analysis focuses on the affordances of the editor, our study addresses how players perceive and evaluate this editor. This includes the customization categories' importance and the perceived sufficiency of the provided options. Harper [16] qualitatively analyzed eight character editors that allow the creation of overweight characters. The analysis considers how these editors allow body size modifications, for example through body-shape sliders, preset body types, and other body-related customization controls. Harper shows that, while such systems may provide control over body shape, they often frame overweight bodies as deviations from an imagined average and restrict the range of possible bodies, particularly through gendered body models and sexual dimorphism. For our study, this illustrates that customization categories should not only be examined in terms of their importance, but also in terms of whether the offered options and degrees of freedom are sufficient. To this end, we also use free-text fields to capture why categories matter to players, which is particularly relevant for categories such as *Body Type* that are implemented through a small number of predefined presets in *BG3*. Complementing this line of work, Pace [33] examines the *WoW* character editor from an

interface-focused, quantitatively-oriented perspective. The work shows that character creation is structured as a sequential series of interdependent decisions. It begins with categories such as race and gender, which can severely limit later choices, such as the selection of hair options based on gender. Using a quantitative analysis of the available options and their combinatorial possibilities per category, Pace showed that the number of available options and possible character combinations differs substantially among customization categories, resulting in a customization space that is hierarchically structured and structurally limited, rather than evenly distributed. These works demonstrate the importance of editorial structure and embedded limitations in character editors. While such expert-driven analyses provide valuable insights into character editor interfaces and their affordances [27], they do not capture how these systems are interpreted or experienced by players. Therefore, we review studies next that incorporate user-centered evaluation approaches.

2.1.2 Studies of Character Editors with Users. In the recent master's thesis by Sheppard [41], *BG3* was considered and a questionnaire administered (n=26). For six categories (of which not all were 1:1 the categories of the *BG3* editor, but inspired by it), participants had to rate whether they tend to customize these. They were also confronted with some selection options from the categories *Race* and *Class* and were asked to indicate which one they would choose. After creating a character, an hour-long think-aloud gameplay session and an interview followed. The results showed that character creation decisions were influenced by both gameplay preferences and aesthetic motives. They also customized all six categories surveyed. However, as category importance was not explicitly asked for and as the character editor includes many more categories, it remains unclear how important the individual categories available in *BG3* are actually rated from the user's point of view. In 2009, Ducheneaut et al. [5] considered character creation in three virtual worlds (*Maple Story* [G7], *WoW* and *Second Life* [G11]). They used an online survey (n=180), aimed at users who had already created at least one character in one of the three virtual worlds. Participants referred to one of the worlds in which they had created at least one character and uploaded character screenshots. In addition, they were asked to rate the perceived importance of each customization category for that specific character editor and to indicate their satisfaction with the number of options available within that category. The customization categories covered in the survey were based on the customization categories available in the corresponding virtual world. Finally, participants provided detailed demographic information about themselves, including their age, height, hair color, and skin color. The results show that hairstyle and hair color were rated as particularly relevant in all three virtual worlds. At the same time, participants stated that they were unsatisfied with the number of options available in these customization categories. Based on our own research of the customization categories available in 2009, *WoW* offered eight [50], *Maple Story* nine, and *Second Life* thirteen categories. In addition, the number of selectable options per category was comparatively limited (each category offered 5–12 options in *WoW* and 3–5 options in *Maple Story* in 2009)⁶. Contemporary character editors have often much higher numbers. For example, the *BG3* editor offers more than 30 categories and has categories with more than 80 (category *Hairstyle*) or even 350 options (category *Skin Color*). Given that Ducheneaut et al. [5]'s studied editors differ in scope and granularity, it remains unclear whether their findings also apply to a contemporary editor, which we investigate in this paper. The results from these authors provide a point of comparison for our work, as they show that users assessed the importance of customization categories differently and that not enough options were provided in 2009. Our study revisits category importance and option sufficiency in the context of a modern editor. While Ducheneaut et al. [5] collected demographic data and related it to participants' character screenshots (see below), individual differences on the category importance assessment were not reported, which we additionally examine.

⁶For *Second Life*, the exact number of available options per category in 2009 could not be determined.

McArthur [28] conducted a qualitative study and assigned participants ($n=24$) to one of two conditions: one group created a Nintendo Wii U *Mii* directly through the *Mii Maker*, while the other group started from a camera-generated *Mii* that could then be edited. Participants also created their *Mii* either alone or in the presence of a familiar person. By combining survey data with observation data, the study shows that character customization is shaped not only by users' self-representational preferences, but also by the concrete creation context, including the starting point of creation, the interface through which options are presented, and the social situation. Although the study did not systematically evaluate the importance and sufficiency of individual customization categories, it identified category-related preferences, such as participants often selecting hairstyles that resembled their real-life hair. The finding that customization practices and category evaluations need to be interpreted in relation to the specific editor and game context in which players encounter them supports our decision to focus on the *BG3* editor as part of a role-playing game, where the created character is not only a visual representation but also central to gameplay, narrative choices, and role-playing for many hours.

2.2 Character Creation

Several studies have explored the relationship between the characteristics of character creators and the traits of the resulting characters [5, 6, 11, 12]. In these studies, participants were either asked to create a character within specific editors and then fill out demographic and personality (e.g., Big Five [13]) questionnaires, or existing characters (e.g., provided through screenshots) were evaluated by third parties based on, for example, attractiveness or likability. The studies used different character editors, ranging from 2D character editors such as WeeWorld⁷ to 3D gaming environments such as *Neverwinter Nights 2* [G9] or *OpenSimulator*⁸. Methodologically, they combined the task of character creation with filling out standardized personality questions. The results show recurring patterns: characters (i.e., decisions on the customization category option level) often reflect aspects of the actual self or an idealized version of it, especially with regard to physical characteristics, attractiveness, or striking visual features such as highly muscular or very slim body types, as well as visually dominant accessories or details [6, 11]. Gender-specific differences also occur, for example in the emphasis on muscular or slim body shapes [5, 6]. Beyond differences related to players' own gender, work shows that the gendered representation of the created character can be meaningful for self-representation and identity exploration: Han and Ho [15] interviewed players about gender exploration and expression in character customization and found that players used customization to conform to or challenge gender norms, symbolize aspects of self-identity, and explore gender through play. Similarly, Björnsson et al. [3] show that the importance of a character's gender and sexual orientation varies across player groups, indicating that character creation choices can be tied to representation, safety, exploration, or narrative intent. These findings suggest that the importance of categories depends not only on gameplay relevance or aesthetic preference, but also on the personal, cultural, or identity-related meanings players attach to them. Our study accounts for this by examining which categories players consider important. The personality traits of the creator can also be reflected in the characters that are created. Relationships to extraversion, neuroticism, and agreeableness were found, while relationships to conscientiousness and openness were less recognizable [12]. Suryapranata et al. [47] investigated whether they are able to create a machine-learning model that is able to predict personality traits based on likeness-values of pre-defined characters (created by the authors based on their own guidelines along general customization categories) that were presented to participants

⁷A social network that ceased to exist in 2017; see <https://en.wikipedia.org/wiki/WeeWorld>, last accessed: 07.07.2026

⁸The website is no longer available.

via an online survey. Their preliminary experiment showed that their model could predict openness, conscientiousness and agreeableness to a certain extent (average accuracy of 66.7%). These findings show that individual differences were already investigated at the level of selection decisions within individual customization categories and connections were found. However, to our knowledge, it has not yet been investigated whether connections can already be found on the customization category level themselves. Our study addresses this gap by examining relationships between individual differences and players' evaluations of customization categories.

Taken together, our study extends prior research. Previous studies have mainly analyzed character editors as designed systems and affordance structures [8, 16, 29, 33], focused on specific aspects, such as gender expression or body representation [3, 15], or related player characteristics to created characters and selected options [5, 6, 11, 12]. In contrast, our study provides a broader, user-centered overview of how players evaluate customization categories in a contemporary RPG. We examine which categories are important, whether their options are sufficient, why categories matter to players, and how these evaluations differ between player groups.

3 Method

We administered an online survey. The language was English and there was no monetary incentive offered for completing it. The study was approved by the ERB of the faculty of Mathematics and Computer Science at Saarland University (No. 25-06-11). The survey was accessible from July 4 to August 19, 2025. Participants were recruited through Reddit⁹, Discord¹⁰ and the official *BG3* forum¹¹. Posts were published in the subreddits *r/BaldursGate3*, *r/BG3*, and *r/SampleSize*. As advertising was not permitted on the official *Baldur's Gate 3* Discord server, smaller community and modding Discord servers were used (for example *BG3 Modding Community*¹²). In addition, seven Twitch streamers (of which one advertised the study) were contacted by email.

3.1 Survey Design

On the start page, we fully disclosed the study's purpose and that it targets people above 18 years of age, who have played *BG3* and created at least one character. After providing consent, (non-mandatory) demographic questions were shown: an inclusive response format was chosen for gender following Spiel et al. [43]. Age was requested in ranges (Under 18; 18-24; 25-34; 35-44; 45-54; 55 or older). People who selected "Under 18" were automatically redirected to the end of the survey. Nationality was recorded in a free text field. Based on the related work indicating that personality traits matter in the context of character creation (and are therefore relevant for **RQ3**), on the same page, the Big Five personality traits [13] were also collected. For this, we used the BFI-10 [36] (using the eleventh item to improve the agreeableness measure), as a fast and validated collection method (given that our survey was long, we decided not to integrate a longer questionnaire).

Afterward, we asked how many characters were created (answer options: 0; 1; 2; 3; ...; 9; more than 9; 0 would lead to an exclusion), the number of game completions (Never; Once; Multiple times), and the estimated playing time (0–1 hours; 2–10 hours; 11–50 hours; 51–100 hours; 101–200 hours; 201–300 hours; more than 300 hours). The scale was guided by published playtime statistics for *BG3*, according to the website *HowLongToBeat*¹³. Based on these values, 51–100 hours was chosen as a typical playthrough; shorter times for beginners, and longer times to represent

⁹<https://www.reddit.com>, last accessed: 07.07.2026

¹⁰<https://discord.com>, last accessed: 07.07.2026

¹¹<https://forums.larian.com>, last accessed: 07.07.2026

¹²Invitation link: <https://discord.com/invite/bg3mods>, last accessed: 07.07.2026

¹³<https://howlongtobeat.com/game/68033>, data from June 2025: Main Story 73h 57m (n=530), Main+Extras 118h 24m (n=2700), Completionist 186h 31m (n=507).

in-depth engagement also through multiple playthroughs. Participants were asked whether they used modifications (mods) (*Yes. I have modded Baldur's Gate 3; No. I have not modded Baldur's Gate 3; I don't remember*). If they did, we asked whether these mods influenced character creation and showed an optional free text field where they could describe how the mods influenced it. All participants were asked to rate two statements on five-point scales¹⁴: *I mostly create characters that are similar to myself* and *I tend to create characters that look similar across different games*.

On the next three pages, we presented 29 customization categories offered in the BG3 character editor (see Table 5)¹⁵. The editor offers even more customization categories, but the following exclusions were made so as not to overwhelm participants:

- **Multi-layered progression decisions.** Certain categories (e.g., *Face, Spells, Subraces*) vary substantially in structure and number of options depending on prior selections (e.g., chosen *Class* or *Race*) and therefore reduce comparability across categories. In addition, we excluded categories tied exclusively to previous decisions and appearing in only one decision path (e.g., *Deities*). However, three categories (*Horn Style, Horn Color, Tail Style*) were retained to include fantasy-specific options, despite depending solely on *Race*.
- **Category refinements that lead to often only subtle changes.** Categories that only adjust the quantitative/material properties of a category (e.g., intensity and quantity for *Freckles*), or color details (e.g., *Horn Tip Color*) were not considered as separate categories.

The presentation sequence of the customization categories was randomized per participant and per page no more than ten categories were shown at the same time. Participants could remind themselves what the category looked like in the game per entry by hovering/clicking on the category name. This showed a GIF showcasing some direct character changes the category enables in the editor. For categories where this was not possible (e.g., *Identity*), the corresponding options tab was shown instead. In addition, we also provide the number of available selection options per category (see also Table 5) in the unmodded BG3 in parentheses (e.g., *Hairstyle* offers 87 options). We denoted the highest number of options, even if certain decisions limit these (e.g., selecting orc as *Race* would limit *Body* from four to two options). Figure 2 exemplifies how both of these look. Per category, participants needed to rate its importance (*How important is the ... category to you personally when creating a character?*) on a seven-point scale with the labels: 1 = Extremely unimportant, 2 = Unimportant, 3 = Somewhat unimportant, 4 = Neither important nor unimportant, 5 = Somewhat important, 6 = Important, 7 = Extremely important. They also needed to rate their perception of the offered selection options (*How did you feel about the number of available options in the ... category?*) within the category on a three-point scale, with the labels: 1 = Too few options, 2 = Sufficient options, 3 = Too many options, and an option to indicate that they do not remember. For participants who used mods, an extra note was visible that they should judge the unmodified game version. To check for careless responses [30], we added one attention check question, which is also visible in Figure 2 (penultimate row).

After having rated all categories, participants had the chance to indicate customization categories that they feel are missing completely from BG3 in an optional free text field (*Were there any customization categories you felt were missing during character creation?*). To learn about reasons for why categories are assessed as important, we aimed for (optional) free text field explanations. We deemed it as too tedious for participants to answer this for all categories that they assessed as at least somewhat important and limited this to their three most important categories. As participants might have provided highest ratings for > 3 categories, we needed to first identify the top three

¹⁴All five-point scales used in the survey had the following labels: 1 = Disagree, 2 = Somewhat disagree, 3 = Neither agree nor disagree, 4 = Somewhat agree, 5 = Agree.

¹⁵Customization categories selection was done based on BG3 version 4.1.1.6848561 (Patch 8).

	How important							Available options			
	Extremely unimportant	Unimportant	Somewhat unimportant	Neither important nor unimportant	Somewhat important	Important	Extremely important	Too few options	Sufficient options	Too many options	Don't remember
Eye Makeup Style (17 options)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Tattoos (45 options)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Vitiligo Pigment (0 to 100)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Heterochromia	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Freckle	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Genitals	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Piercings (1 to 10)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Eye Colour (1 to 100 options)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Attention Test (Select "Unimportant", "Too few options")	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Tattoos Colour (62 options)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Fig. 2. Excerpt of the customization rating page, showing the hover interaction on the category Tattoos. Tattoo example screenshot taken from the character editor of *Baldur's Gate 3* by Larian Studios [G1].

specifically by presenting the question *Imagine that only three of these customization categories could remain in the character editor and the others would be removed. Which ones would you choose to keep?* and presented all categories having received the highest rating with check boxes, at which point they needed to select three. After having identified the most important categories, either through the above mentioned process or automatically (if only ≤ 3 categories received the highest rating before), and the provided highest rating was above four on the seven-point scale, we showed an optional free text field with the question *Why is the customization category ... especially important to you?* for every top category. We then presented four statements, each to be rated on a five-point scale: All started with *The Baldur's Gate 3 character editor was...* and ended with: *...the best character editor I have ever used; ...overall good; ...sufficient to create the character I wanted; and ...lacking in many aspects.* An optional free text question, asking for reasons and games with better editors, appeared if participants (somewhat) disagreed to the first statement. An optional free text field was also shown if participants (somewhat) agreed to the last statement, asking them what they felt was missing. On the last page, participants could (optionally) provide general comments. The survey then also allowed them to upload one *BG3* save file, which we planned to additionally analyze. However, as only 29 usable save games were uploaded, we excluded this for this paper.

3.2 Analysis Approach

We briefly illustrate our qualitative and quantitative analysis approach in this section.

3.2.1 Customization Category Classification. Considering the 29 customization categories, a general difference is whether they have an impact on the game (e.g., on the character's available actions, impact on dice throws or on dialog) or are just purely cosmetic. The authors analyzed every chosen category and discussed potential effects categories have on the game. To this end, both authors (for their background see also next section) experimented within the game and complementary to this, one author also checked various online sources. Overall, five of the 29 categories were identified to have a game impact: *Ability Points*, *Background*, *Identity*, *Class* and *Race*. Of these, only the last two categories also have an impact on character appearance in the editor.

3.2.2 Qualitative Content Analysis. We conducted a conventional qualitative content analysis [19, 44] for the free text answers. To increase the coding quality, both authors considered all answers and each author created an individual code book for every question based on the inspected answers/questions first. Experience-wise, one author played *BG3* for more than 400 hours, created more than 5 characters, and finished the game three times. The other author has played *BG3* for more than 30 hours (with only one character created, game not finished). In addition, both authors had played more than 20 RPGs and each has several years of tabletop RPG experience (both as a game master and a player), including various D&D rule versions. The two resulting code books per question were then merged into one through discussion. Then, the authors independently coded all answers with the corresponding merged code books again. Afterwards the results were compared. Where deviations were identified, a discussion followed to identify the reasons. As the reasons were mainly based on interpretation nuances or because something was overlooked, all deviations could be solved easily. Thus, the final code assignments are the results of a 100% agreement between the authors. Together, they identified larger themes were reasonable [26].

3.2.3 Statistical Analysis. We primarily followed the recommendations of Field et al. [10]. The median (Md) is reported in the descriptive analyses. As it is reported in similar work as well, we also show mean (M) values, although this is controversial (because of actual ordinal data, cf. [10]; see also¹⁶). We also denote the percentage of participants that agreed to a statement (Agreement Rate (AR))¹⁷. We treat the Big Five traits per participant as low or high (which is not uncommon, e.g., [51]) based on the participant's result on every subscale (low: ≤ 2 ; high: ≥ 4).

For statistical inference, Shapiro-Wilk tests were first performed to verify the assumption of normal distribution. Since the distributions in all categories deviated significantly from a normal distribution, non-parametric methods were used instead. For each customization category, it was checked whether the importance ratings deviated significantly from the center of the seven-point scale by using the Wilcoxon signed-rank test. For comparisons between two groups the Wilcoxon rank sum test, and for comparisons between more than two groups the Kruskal-Wallis test was used. Significant differences identified by the Kruskal-Wallis test were then calculated using post-hoc Dunn tests and Holm correction. Due to the exploratory nature of our research, multiple tests are controlled using the false discovery rate (FDR) according to Benjamini and Hochberg [2]. This method is less strict than the Bonferroni correction and thus reduces the probability of falsely rejecting potentially relevant effects [10]. In general, the test statistics reported are the FDR-corrected p-value; effect sizes are reported with Cohen's r ($.1 = \text{small}$, $.3 = \text{medium}$, $.5 = \text{large}$ [10]).

3.2.4 Exploratory Factor Analysis (EFA). An EFA was conducted based on the importance ratings of the customization categories. The following requirements were checked according to [9]:

- (1) All items must have a correlation of at least .3 but no more than .9 with another item.
- (2) Participant count (N) should exceed the item count (x) by a factor of at least 3 ($N > x \cdot 3$).
- (3) Values from the Kaiser-Meyer-Olkin calculation (KMO) between .8 and 1 are considered sufficient. Values below .6 are insufficient [21].
- (4) The Bartlett test should show significance.

To calculate the factors to be extracted, we use Horn's parallel analysis method [18]. This calculation method is reliable because it compares the eigenvalues of extracted factors with simulated factors at the same position.

This reduces the influence of random data variations on factor extraction [4]. We applied an Oblimin rotation method, because participants were expected to show high relevance ratings

¹⁶<https://stats.stackexchange.com/questions/67551/calculate-mean-of-ordinal-variable>, last accessed: 07.07.2026

¹⁷On the seven-point scale, approval is rated from five onward, and on the five-point scale from four onward.

across multiple factor dimensions, rather than loading on a single exclusive factor, similar to the assumptions in Schuck et al. [39]. To determine the factor structure, all items were evaluated based on their factor loadings. Items were assigned to the factor on which they had their highest loading. A value of at least .4 was used as the minimum criterion for a substantial loading, in accordance with [45]. Items that did not meet this criterion or showed significant cross-loadings were excluded from further interpretation. Cronbach's α was calculated to assess the internal consistency of the factors. Higher values indicate greater internal consistency, with values above .7 considered acceptable and values above .8 considered good [32]. Finally, the clusters formed were named and interpreted by the authors after the content of the associated items had been checked.

4 Participants

191 participants completed the survey but 20 failed the attention check and were excluded; no other exclusion criteria were met (final sample: 171; gender: 58x women, 92x men, 19x non-binary, 6x without specification and five participants selected more than one checkbox; age: 39x 18–24, 84x 25–34, 31x 35–44, 13x 45–54, 4x 55+). Most participants (67x) came from Germany and the USA (35x), with the rest spread across the world, but with a Western bias (151 of 160 who answered this question). The Big Five traits distribution (low (l)/high (h)) was uneven (Extraversion l: 80, h: 24, $M = 2.49$; Agreeableness l: 10, h: 70, $M = 3.61$; Conscientiousness l: 19, h: 46, $M = 3.21$; Neuroticism l: 38, h: 73, $M = 3.32$; Openness l: 9, h: 94, $M = 3.81$). The sample showed a high level of experience with *BG3* (playing time: 1x 0–1h, 1x 2–10h, 10x 11–50h, 17x 51–100h, 23x 201–300h, 85x 300+h; number of created characters: 16x 1, 19x 2, 20x 3, 15x 4, 17x 5, 12x 6, 2x 7, 8x 8, 1x 9, 61x 10+; completed the game: 41x never, 42x once, 88x multiple times).

We defined participants as power users if they reported (1) at least 200 hours of playtime and (2) having created at least six characters. The playtime threshold was chosen to capture substantially above-average engagement with *BG3*, based on the playtime statistics described above. The second criterion reflects at least an average use of character creation, ensuring that power users not only accumulated playtime but also used the editor to a larger extent. Since no external reference values for the average number of characters created per player were available, this threshold was derived from our sample: The mean number of created characters was 6.01, with the open-ended category “10+” coded as 10; we therefore used six characters as a conservative lower bound. Based on these criteria, 76 participants were classified as power users and 95 as non-power users. 75 participants stated that they had expanded the character editor through mods, which are analyzed in a later section. The evaluation of the statements on character creation showed that participants tended not to create characters based on themselves ($M = 2.03$, $Md = 2$, $AR = 18.1\%$). In contrast, there was a more balanced response when it came to creating the same characters across different games ($M = 2.84$, $Md = 3$, $AR = 47.4\%$).

5 Results

Findings are presented as a headline together with a number (R=result).

5.1 Overall Perception of the *BG3* Character Editor

We first focus on the questions in respect to the character editor in general.

R01: The character editor is perceived well and allows the creation of desired characters.

The statement that the character editor is good overall received a high level of agreement on the 5-point scale ($M = 4.55$, $Md = 5$, $AR = 93.0\%$) as well as the statement that the editor is sufficient to create the desired character ($M = 4.26$, $Md = 5$, $AR = 84.2\%$).

R02: The character editor offers potential for improvement, especially in respect to customization categories and options within these.

The sample was indifferent whether the BG3 character editor was the best editor they had ever used ($M = 3.07$, $Md = 3$, $AR = 46.8\%$). 48 of the 62 participants who (somewhat) disagreed to this statement provided answers to the question what was missing and/or provided game names with better editors. 30/48 provided game(s) with, in their opinion, better character editors.

Better editors. 25 different games with subjectively better editors were mentioned; those receiving at least three mentions are: *The Sims* (here in particular *The Sims 4* [G13]) (9x), *Elden Ring* [G5] (6x), and *The Elder Scrolls V: Skyrim* [G12] (5x, three times explicitly mentioned with mod usage), *Cyberpunk 2077* [G3] (3x), *Black Desert Online* [G2] (3x) and *Dragon Age: The Veilguard* [G4] (3x).

Shortcomings of the editor. 39/48 participants answered the other part of the question. Most wanted finer control within existing categories (26x), more options in existing categories (10x), or more appealing options (4x). While overall twelve existing categories were mentioned across participants, criticism (e.g., “*The lack of body type options is the biggest issue for me. Also the choice in faces is really quite sparse*”), focused primarily on *Face/Head* (19x) and *Body* (13x), while missing categories were mentioned only three times (we will elaborate on these in R07). Overall, 22 participants (19% of the sample) felt character creation was too limited in general and/or that they cannot create their intended character with the current set of options (e.g., “[...] *you cannot build a female buffed character*”). Participants generally disagreed that the editor lacked many aspects ($M = 2.45$, $Md = 2$, $AR = 25.1\%$), suggesting that the above issues were not widespread. 25/43 participants provided comments to the respective free text question, mainly criticizing *Face/Head* (15x) and *Body* (11x). Seven feature suggestions were also submitted (see R03).

Mod usage. Mod¹⁸ usage can give insights into editor’s improvement potential. Of the 75 participants who reported using character creation mods, all described their changes in a free-text field, though seven responses were unclear. Among the remaining 68, 65 used mods adding options to existing categories (will be elaborated on in R09), 23 used mods introducing new categories (see R07), and only seven used mods adding new features (see R03).

R03: Additions, Removing Restrictions, UI/UX-related Improvements and Subjective Criticism-related Changes are themes for which features were suggested.

Suggestions beyond new or expanded categories were submitted in free-text responses. Although only 24 participants provided such, often mentioned just once, we included all 26 in Table 1 as a resource for developers and researchers. We found four overarching themes: *Additions* represents features that are not yet present. *Removing Restrictions* shows that some participants are not content with aesthetic/mechanical restrictions that are currently implemented, i.e., functionality is (mainly) already available, but restricted to previous decisions (e.g., selected *Class*). *UI/UX-related Improvements* covers elements that would improve the understanding/handling of the editor and *Subjective Criticism-related Changes* covers subjective issues with certain (game) design decisions.

Section Takeaway. The BG3 character editor is neither perfect nor the best editor available. Considering R01 and that criticism is focused mainly on specific categories (primarily lack of depth for *Body* and *Face/Head*), instead of general usability problems, we deemed the editor as suitable for analyzing how players prioritize and evaluate customization categories. Concerning the free text answers, we learned that the issues are primarily concerning the currently offered options within the customization categories and too coarse-grained categories, less so missing features.

¹⁸The number of BG3 mods is large, e.g., <https://www.nexusmods.com> alone has 17,651 (as of 07.07.2026) BG3 entries.

Table 1. List of feature suggestions beyond adding/extending customization categories in particular, sorted by themes (number in parentheses indicate how many participants contributed at least one suggestion to the corresponding theme). (3x) or (2x) indicates that the suggestion was made by more than one participant.

Feature Suggestions	Feature Suggestions
Additions (11/24) Ability to alter the appearance of the origin characters (3x) Visibility of camp clothes in the editor Fine-grained options for coloring <i>Tattoos</i> (e.g., color transitions) Possibility to select more than one <i>Background</i> Option to adjust healthiness of skin appearance (e.g., patchy skin) Morph option for different visual features Allowing to drag different <i>Body</i> parts in the editor Possibility to make <i>Tattoos</i> glossy/metallic Possibility to adjust hair length for existing <i>Hairstyles</i>	Removing Restrictions (9/24) Unlock every <i>Face</i> option for every character (2x) Unlock <i>Tails</i> for every character (2x) Unlock <i>Horns</i> for every character Unlock all color options for all color-based categories Unlock <i>Deity</i> selection for every character Removal of <i>Spell</i> restrictions Removal of <i>Skill</i> restrictions Removal of <i>Body</i> and <i>Clothing</i> connection Removal of <i>Hair Color</i> and <i>Facial Hair Color</i> connection Removal of <i>Hair Color</i> and <i>Eyebrow Color</i> connection
UI/UX-related Improvements (3/24) Providing more explanations in the editor Free camera movement around character in the editor Direct selection of elements, instead of clicking through a selector Better information arrangement (some categories are too hidden)	Subjective Criticism-related Changes (3/24) Bonus ability points based on race Better sexual dimorphism Editor assumes white people as default

5.2 Perception of the Customization Categories

We present results in respect to participants' general perception of the customization categories.

R04: The majority of customization categories are important for character creation.

Figure 3 shows the aggregated importance ratings for the categories. 23/29 categories had a mean larger than the (neutral) midpoint of the 7-point scale and 19/23 deviated significantly from it (with 16 having a medium to large effect size). This shows that participants generally consider a wider range of customization categories to be relevant for creating their character.

R05: The most important categories are *Class*, *Race*, *Hairstyle*, *Ability Points*, *Body Type* and *Hair Color*, while *Genitals* and *Vitiligo Pigmentation* were considered the least important.

Most important. The explicit top three selections (see Figure 4) reveal four dominant customization categories: *Class*, *Race*, *Hairstyle* and *Ability Points*. The *Class* category stands out in particular with 122 appearances (71.3%), while the other three were each selected by more than a third of the sample. *Body Type* and *Hair Color* follow at a considerable distance. Unsurprisingly (method-wise), this pattern is reflected in Figure 3: the six categories have high mean values and large effect sizes (of at least .72). While the scale rating allows for a differentiated assessment of all categories, the top three reveal particularly dominant categories. Considering the selection of the three most important categories, the following nine categories were never among them: *Horn Color*, *Hair Graying*, *Tail Style*, *Scarring*, *Piercings*, *Eye Makeup Style*, *Eye Makeup Color*, *Lip Tint*, *Hair Highlights*.

Least important. *Genitals* and *Vitiligo Pigmentation* were the only elements that deviated significantly from the midpoint of the scale (with medium to large effect sizes) and had a mean value below four (see Figure 3). These categories are less relevant for the sample, but we have to note that some participants found them relevant (seen by mean, median and AR values and that *Genitals* and *Vitiligo Pigmentation* were mentioned twice/once as one of the most important categories).

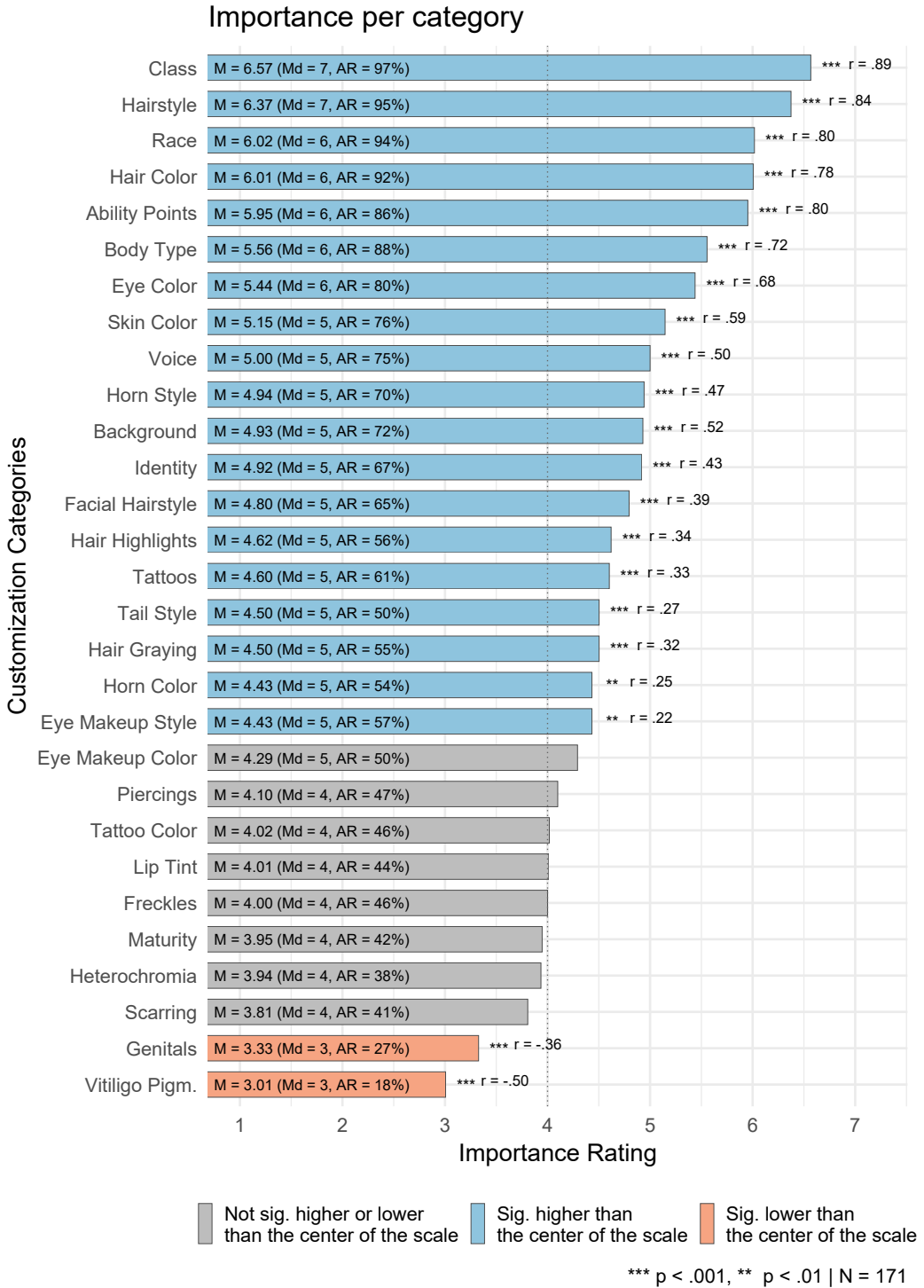


Fig. 3. Importance of the 29 customization categories compared to the scale midpoint (4 = neutral). Mean (M), median (Md), effect size (r), agreement rate (AR), and significant deviation from the scale midpoint.

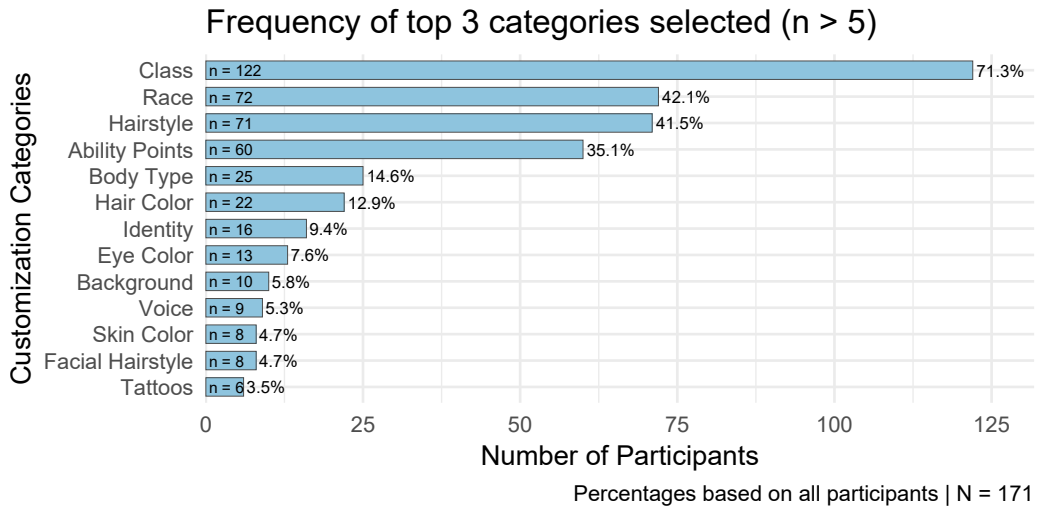


Fig. 4. Frequency of customization categories selected as top three by at least five participants (n).

Table 2. Reasons why a category was seen as important and how many participants indicated these.

Reason	Description	Freq.	Reason	Description	Freq.
	Creating Imagined Character			Direct Game Impact	
Relevancy for character	Category allows creating the intended character visually or personality-wise	90	Gameplay	Category has an impact on game aspects or playstyle (e.g., dialog or mechanics)	138
Role-playing	Category allows engaging in role-playing and staying in character in-game	53	Replayability	Decisions within the category add to the game's replayability	27
Unique touches	Category allows adding unique touches to the character/allows individualization	39	Appearance less relevant	Participants stated that a category has a larger impact than being visual-only	19
Diversity	Category allows creating diverse characters	24	Adding immersion	Category adds to the immersive experience	9
Fiction	Category allows creating characters that cannot (easily) be created in reality	15		Aesthetic Relevancy	
Creative freedom	Category is important because it offers experimentation	10	Aesthetic reasons	Category plays a significant role for the look of the character	46
Recreate self	Category allows creating characters similar to oneself or parts of oneself	4	Visibility	Category is clearly visible	30
			Relevancy for other elements	Category is relevant in combination with other elements	23
			Exposure time	Participants stated that a category is important because of the game's length	9

R06: *Creating Imagined Character*, *Direct Game Impact* and *Aesthetic Relevancy* are themes for why categories were selected as important.

147 participants contributed reasons why a category is important to them. 42 participants mentioned something related to personal preference (e.g., “*Hair is cool, lmao*”). 17 other reasons are listed in Table 2 (and their distribution in Table A1). We identified three overarching themes:

- **Creating Imagined Character** consists of reasons that represent that participants have a clear notion on what they want to create and valued categories that helped them to realize this vision. Frequency-wise, the second most frequent reason represents this specifically (e.g., “*It's the key feature for your characters identity*”); followed by role-playing aspects, i.e., that a category helps participants make decisions in game situations (e.g., “*It influences through which*

Table 3. Themes (why a category is important) × categories. Categories without any answer: *Horn Color, Hair Graying, Tail Style, Scarring, Piercings, Eye Makeup Style, Eye Makeup Color, Lip Tint, Hair Highlights, Vitiligo Pigmentation*. **N** = No. of participants who were asked question for the category; **n** = No. of responders. **X** = No. of participants who contribute to the theme at least once. Cells show the number of participants that responded to at least one code within the theme for that category. Bold highlighting indicates elements that contribute at least 10% to the theme. [Table A1](#) shows distribution per code (instead of themes).

Theme n/N	X	Class 103/122	Race 58/72	Hairstyle 64/71	Ability Points 50/60	Hair Color 17/22	Body Type 19/25	Eye Color 12/13	Identity 11/16	Background 9/10	Voice 9/9	Skin Color 8/8	Tattoos 6/6	Facial Hairstyle 6/8	Horn Style 2/3	Maturity 2/3	Freckles 2/2	Genitals 2/2	Vitiligo P. 0/1
Creating Imagined Character	192	34	33	36	21	7	17	5	9	7	6	5	3	3	2	2	0	2	0
Direct Game Impact	161	75	34	1	39	1	2	0	2	3	2	2	0	0	0	0	0	0	0
Aesthetic Importance	81	2	12	33	6	8	5	4	0	0	3	4	2	2	0	0	0	0	0

lens I see the character”). Other reasons are that a category helped to add unique touches (e.g., “[...] it’s a good way of making different characters stand out a bit more”), that it is important for diversity/representation reasons (e.g., “As an enby I want to feel represented”), or that it allows experiments and offers creative freedom (e.g., “[the category] provides a lot of freedom in creating different builds”). Some participants also want to re-create elements that they have in real life (e.g., “I just really like looking like my irl self to project me into the game for a more immersive feeling”); while others explicitly mentioned to create something that is only possible in a fictional world (e.g., “You can play characters that are outside what we know from the real world”).

- **Direct Game Impact** consists of reasons with a strong impact on the game itself (e.g., “It has an actual effect during the gameplay”) or the game’s replayability (e.g., “The different racial benefits also add to replay value”). Some participants also highlighted the increased importance of a category in comparison to “just” appearance-related aspects (e.g., “I care more about gameplay than looks I guess”). Finally, we see immersion reasons (e.g., “It’s very important for immersion”) as part of this theme, as this might keep players longer engaged within the game.
- **Aesthetic Relevancy** consists of aesthetic reasons. The most-frequent reason underlines that a category plays a significant role for the whole appearance of the character (e.g., “Hairstyles change the whole disposition of the character”); is prominently visible (e.g., “[...] the most prominent feature in top down view”); are of significant relevance for other categories in the editor (e.g., “The race has an impact on a lot of the cosmetic options”) and because of the long impact the decision has (e.g., “A voice I like is more pleasant than one I don’t. Especially in such a long game”).

Considering the category distribution to these themes (see [Table 3](#)), categories that have an actual impact on the game relate strongly to *Direct Game Impact*, but also to *Playing Imagined Character*. It is interesting to see that *Hairstyle* is not just relevant for *Aesthetic Importance*, but also for *Creating Imagined Character*. Although *Hairstyle* was not the most important category (which was *Class*) (see [R05](#)), in both themes, it was the dominant category.

R07: The BG3 character editor does not offer enough customization categories.

69 participants mentioned overall 33 categories they felt missing during character creation. Although many categories were only named once or twice, we deemed it reasonable to provide

Table 4. Listing of customization categories that participants felt lacking in the editor. For overview purposes, clustered into groups. Number in parentheses shows how many individual participants added to this cluster.

New customization category	Freq.	New customization category	Freq.
Body-related (26/69)		Face/Head-related (25/69)	
Body proportions in general	10	Head proportion in general	21
Height	9	Eye	4
Weight	9	Nose	4
Muscle amount	7	Mouth	4
Breast size	2	Ear	2
		Chin	2
		Cheek	2
		Jaw	1
		Forehead	1
Aesthetic details (16/69)		Uniqueness (16/69)	
Body tattoos	12	Starting gear	6
Nail polish	2	Gait	2
Body hair	2	Posture	2
Body scars	1	Wings	2
Nail types	1	Facial expression	1
Accessories	1	Adding artificial limbs	1
Arm hair	1	Character portrait pose	1
Leg hair	1	Quirks (with impact on conversations)	1
Eyebrow color	1	Negative traits	1
Glasses	1		

the full list, as a resource for developers (see Table 4). We clustered these into groups for overview purposes only (i.e., one row could also be part of another cluster). This shows, as also seen in R02, that many participants have issues with *Body* and *Face/Head*. Based on the free text answers on mod usage (see R02), we found that these were used to add these new categories: changing starting clothing (mentioned by 9 participants), changing head form (9x), accessories selection (3x) and changing the eye form (4x), while the following aspects were only mentioned once each: changing clothing colors, nail style, nail colors, wings, teeth colors and body tattoos.

Section Takeaway. Although not every customization category of *BG3* is important for everyone, all offered categories appealed to someone, showing that offering many categories is reasonable and that even the modern editor of *BG3* does not offer enough. In addition, there are categories that are clearly important for many participants and should be present in all editors. There are diverse reasons why a category is seen as important, yet we were able to find overarching themes showing that representation, aesthetic and gameplay impact matter.

5.3 Perception of the Customization Categories Options

In this section, we focus on the participants' perception of the offered options within categories.

R08: Customization categories offering many selection options do not overwhelm users.

Table 5 shows the customization categories together with the assessment of the offered selection options per category. Only a small percentage rated the available options as too many, showing that offering many options appears to be unproblematic. Even the category with the most options—*Skin Color*: 350—received a too many rating from only 22 participants (12.9% of the sample).

Table 5. *BG3*'s customization categories used in our survey (with indication whether these have a game impact), their rank (1–29) in Figure 3 (i.e., lower = higher importance rating) and their max. available options in the unmodified *BG3*. Per category, we show how many participants (N; excluding participants answering “Don’t remember”) answered how they perceived the number of options (too few, sufficient, too many). We sorted the table based on the rounded perceived option insufficiency percentage (POIP) ($= n \text{ too few} / N$).

Category	Rank	# Options	N	n too few	n suff.	n too many	POIP
Body Type	6	4	168	125	42	1	74
Hairstyle	2	87	169	106	62	1	63
Tail Style	16	4	140	78	61	1	56
Facial Hairstyle	13	19	151	79	72	0	52
Voice	9	8	167	89	76	2	53
Piercings	21	16	159	85	69	5	53
Tattoos	15	45	165	86	74	5	52
Horn Style	10	15	152	67	83	2	44
Background (game impact only)	11	11	166	65	96	5	39
Scarring	27	10	159	62	90	7	39
Eye Makeup Style	19	17	157	56	98	3	36
Race (game + aesthetic impact)	3	11	171	60	107	4	35
Hair Color	4	125	170	44	119	7	26
Genitals	28	6	159	29	120	10	18
Class (game + aesthetic impact)	1	12	171	27	141	3	16
Horn Color	18	99	149	22	121	6	15
Eye Color	7	145	169	26	129	14	15
Maturity	25	0–100 slider	162	22	137	3	14
Hair Highlights	14	163 + 0–100 slider	167	23	132	12	14
Ability Points (game impact only)	5	27 points to distribute on 6 9-pt. scales	169	18	150	1	11
Tattoo Color	15	62 + 0–100 slider	161	17	133	11	11
Freckles	24	2x0–100 slider	164	16	142	6	10
Skin Color	8	350	168	16	130	22	10
Eye Makeup Color	19	94 + 3x0–100 slider	155	14	135	6	9
Vitiligo Pigmentation	29	0–100 slider	150	12	129	9	8
Lip Tint	23	77 + 3x0–100 slider	158	13	133	12	8
Identity (game impact only)	12	3	162	13	137	12	8
Hair Graying	17	143 + 0–100 slider	161	12	143	6	7
Heterochromia	26	145	164	6	146	12	4

R09: Several customization categories need more selection options.

Table 5 also shows that many participants need more selection options: For seven categories, more than 50% of the sample found that these lack options. Considering categories for which at least 25% want more options, this number rises to 13. Within these, there are categories that already offer many options (*Hairstyle*: 87, *Hair Color*: 125, or *Tattoos*: 45), showing that even these numbers are not enough for many participants. Most participants who used mods (see R02) added selection options to existing customization categories. From the 24 corresponding codes, ten were mentioned at least seven times: *Hairstyle* (52/68 participants used mods for this), *Face/Head* (26x), *Classes/Subclasses* (21x), *Races/Subraces* (17x), *Horn Styles* (12x), *Eye Color* (11), *Skin Color/Skin Texture* (10x), *Piercings*, *Makeup* (7x), *Hair Color* (7x) (the rest was mentioned below five times each).

Section Takeaway. The results reveal that the *BG3* editor does not offer enough options in several categories, although some of these already offer many options. Given that the majority of participants do not think the selection options within categories are already too much, it appears to be unproblematic to provide additional options to choose from.

Table 6. Factor structure of the categories according to the EFA. Excluded customization categories: *Background, Race, Identity, Voice, Maturity, Hair Highlights, Hair Graying, Facial Hairstyle, Horn Color*.

Factor	Customization categories	Cronbach's α
F1: Detail Aesthetics	<i>Scarring, Freckles, Vitiligo Pigmentation, Tattoos, Tattoo Color, Piercings, Eye Color, Heterochromia, Eye Makeup Style, Eye Makeup Color, Lip Tint</i>	.91
F2: Basic Appearance	<i>Body Type, Skin Color, Hairstyle, Hair Color</i>	.76
F3: Body Features	<i>Genitals, Horn Style, Tail Style</i>	.75
F4: Mechanics	<i>Class, Ability Points</i>	.66

5.4 EFA Results and Individual Differences

Gender, age, personality traits, and *BG3* usage were examined as discriminating factors for customization category importance ratings and character editor statements. For the character editor ratings, no significant differences across factors were found for either the overall evaluation (see **R01**) or the statement that it was the best editor ever used (see **R02**; all $p > .05$).

R10: The customization categories can be clustered into four factors: *Detail Aesthetics, Basic Appearance, Body Features and Mechanics*.

Conditions. An EFA was conducted to identify potential further structures within the customization categories: The customization category *Background* was excluded, as it did not show sufficient correlation ($r < .3$) with any other category (see method section); for the other 28 categories, the correlation requirements were met. The required sample size ($171 > 84 = 3 \cdot 28$) was met. Eight categories were excluded (see **Table 6** caption) due to low factor loadings or relevant cross-loadings, resulting in 20 customization categories retained for factor interpretation. Horn's parallel analysis suggested a four-factor solution. The final solution was obtained with Oblimin rotation and explained 61.3% of the variance. The resulting factor structure is shown in **Table 6**. The KMO value was high at .87, and the Bartlett test showed a factorizable correlation matrix ($p < .001$). The internal consistency of the factors was in the acceptable to high range ($\alpha = .66 - .91$). F4 was retained despite its lower internal consistency ($\alpha = .66$), because it contains only two items and is clearly interpretable as the only mechanics-related factor. Since Cronbach's α is limited for two-item scales, we additionally report the Spearman-Brown reliability estimate, which is recommended for two-item scales [7]. The coefficient was slightly higher ($\rho_{SB} = .69$), supporting the retention of the factor while still warranting cautious interpretation.

Factors. F1 ("*Detail Aesthetics*") groups visual details that modify or decorate an already defined appearance, such as *Scar* or *Tattoo*. F2 ("*Basic Appearance*") includes broad appearance categories like *Body Type* and *Hairstyle* that shape the character's overall look. F3 ("*Body Features*") includes specific body features that are, to our knowledge, not commonly adjustable in other editors. F4 consists of categories with a game mechanics impact ("*Mechanics*"). The factors indicate that game mechanics are distinct from visual features, and that functional character development and aesthetic design are perceived separately in editors. Instead of one aesthetic cluster as expected, we found three, suggesting participants perceive specific aesthetic options differently.

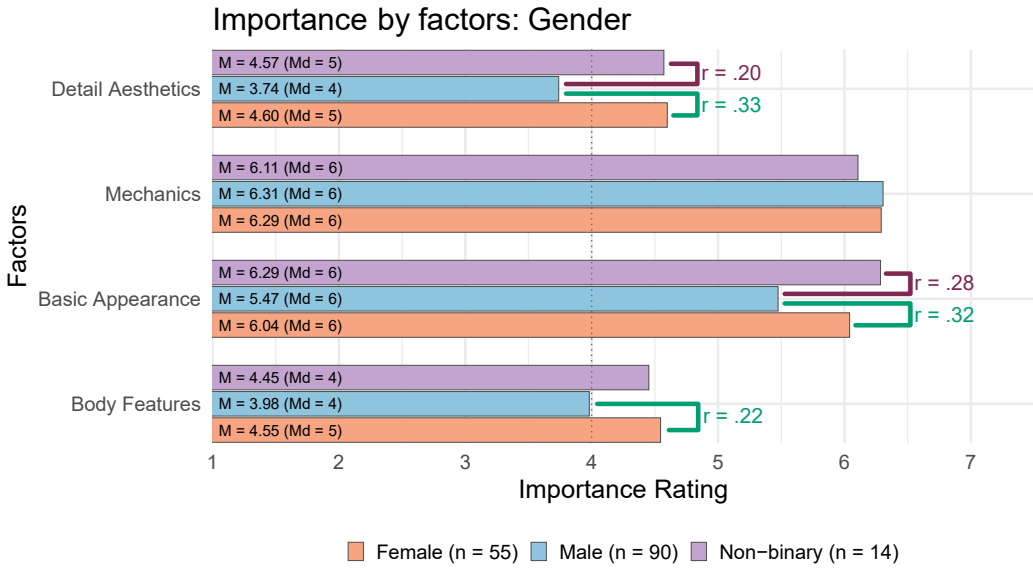


Fig. 5. Comparison of the importance of the four factors between genders (group-test non-significant for *Mechanics* only). Number of participants (n), mean (M), median (Md), significant difference for post-hoc tests (connecting lines between groups) and effect size (r).

R11: Compared to female participants, male participants rated the character editor as more sufficient for creating desired characters.

A significant gender¹⁹-related difference ($H(2) = 12.54, p = .004, N = 159$) was observed regarding whether the *BG3* editor is sufficient for creating desired characters (cf. **R01**): male ($n = 90, M = 4.53, Md = 5$) reported higher agreement than female ($n = 55, M = 3.98, Md = 4$; $p = .002, r = .27$) participants with a small to medium effect size. No sig. differences were found between non-binary ($n = 14, M = 4.00, Md = 4.5$)/female and non-binary/male participants. No sig. differences for this statement were found for age, power usage, or personality traits (all $p > .05$).

R12: Female and non-binary participants see more deficits in the character editor.

For the statement of whether the editor is lacking in many aspects (cf. **R02**), a significant gender-related difference was found ($H(2) = 18.41, p < .001$): female ($M = 2.89, Md = 3$) and non-binary participants ($M = 2.92, Md = 3$) agreed more with it than male participants ($M = 2.01, Md = 2$, female/male: $p < .001, r = .31$; non-binary/male: $p = .02, r = .2$) with a small to medium effect size. No sig. differences for this statement were found for age, power usage, or personality traits.

R13: Female and non-binary participants rate aesthetic-related aspects as more important than male participants.

For the customization category importance ratings (cf. **R04** and **R05**), significant gender-related differences were found only for the three aesthetic clusters with small to medium effect sizes (see **Figure 5**): female participants rated all three clusters, and non-binary participants two clusters, as significantly more important than male participants. At the category level (see **Figure A1**), female

¹⁹For gender comparisons we excluded participants that selected more than one gender, due to the small group size (5x).

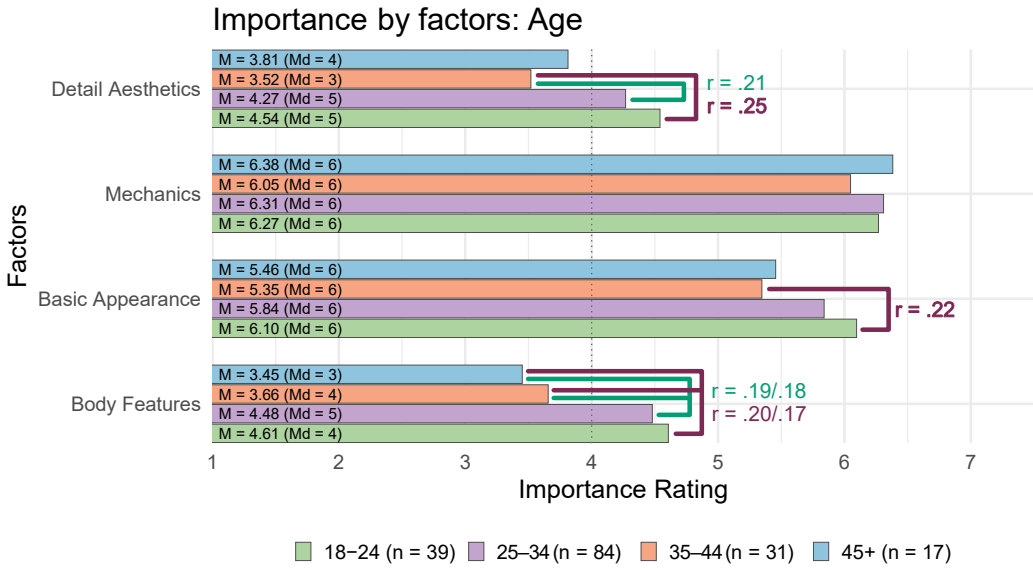


Fig. 6. Comparison of the importance of the four factors between age groups (group-test non-significant for *Mechanics* only). Number of participants (n), mean (M), median (Md), significant difference for post-hoc tests (connecting lines between groups) and effect size (r).

participants rated 13 and non-binary participants nine categories as significantly more important than male participants. *Identity* is the only category with a marginal gameplay impact²⁰.

R14: Younger participants rate certain aesthetic aspects as more important than older participants.

Significant age-related differences were found only for the three aesthetic clusters with small to medium effect sizes (Figure 6). At the category level (Figure A2), differences were observed for *Hairstyle*, *Horn Style*, *Tattoos*, *Tail Style*, *Heterochromia*. Younger participants tended to rate these categories as more important, although this pattern was not consistent across all pairwise comparisons. The lack of sig. differences involving the oldest group may be due to its smaller size.

R15: Participants scoring high on neuroticism rate aesthetic-related aspects as more important than those scoring low.

For the Big Five traits, significant differences in customization importance ratings were found only for neuroticism and openness. No significant differences were observed for extraversion, agreeableness, or conscientiousness at either the category or factor level. Although openness showed nine significant differences, the highly unbalanced group sizes ($n_{low} = 9$; $n_{high} = 94$) limit interpretability, and no conclusions are drawn from this finding. For neuroticism, significant differences with effect sizes ranging from small-to-medium to medium were found for all factors except *Mechanics* (see Figure 7), with participants scoring high on neuroticism rating categories as more important. At the category level (Figure A3), 21 of 29 customization categories were rated as significantly more important by participants scoring high on neuroticism. Among the gameplay-related categories, only *Identity* showed a significant difference.

²⁰ A small number of dialogs use a different forms of address based on selections in this category.

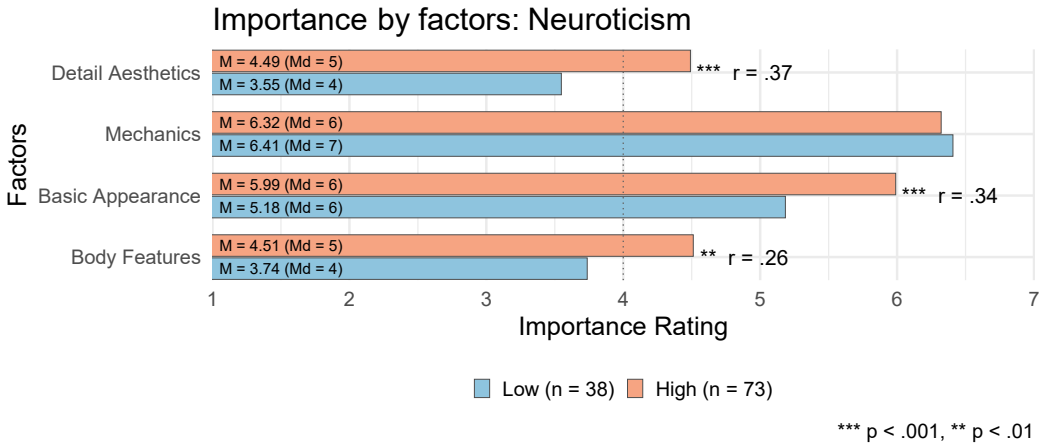


Fig. 7. Comparison of the importance ratings of the four factors between low/high neuroticism. Mean (M), median (Md), effect size (r), and significant difference.

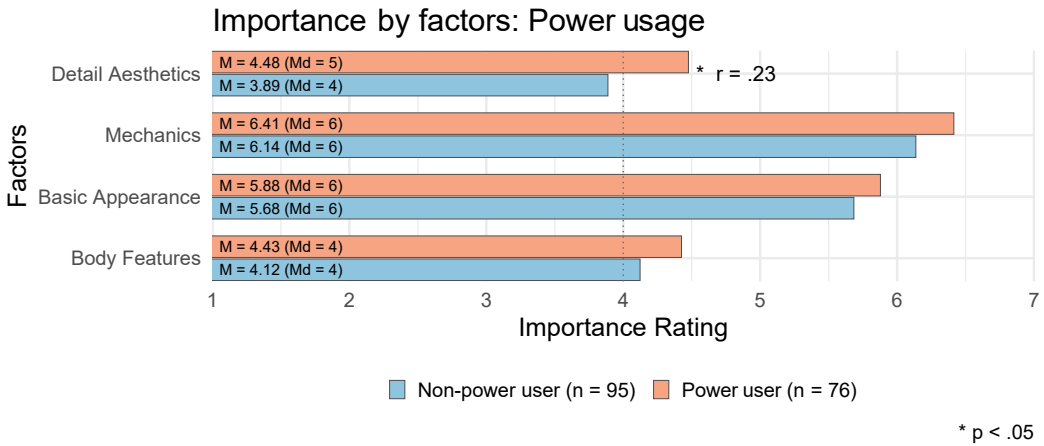


Fig. 8. Comparison of the importance ratings for the four factors between power and non-power users. Mean (M), median (Md), effect size (r), and significant difference (*).

R16: Power users rate aesthetic details as more important.

For power users, a significant difference was found only for the *Detail Aesthetics* factor with a small to medium effect size (see Figure 8). At the category level (Figure A4), power users rated six categories as more important than non-power users; five belonged to this factor, while one had been excluded from the EFA.

Section Takeaway. Across the discriminating factors systematic differences in customization preferences emerged, particularly for aesthetic-related categories. However, these findings reflect group differences in perceived importance rather than causal effects, and are generally of small to medium magnitude. Overall, participants vary in how strongly they value cosmetic customization options, while gameplay-relevant categories show little or no consistent variation.

6 Discussion

We will discuss key points and provide design suggestions that we derive from the results.

6.1 Overall Perception and Shortcomings of the Character Editor

Regarding **RQ 1**, the overall positive evaluation of the editor (**R01**) coexisted with specific shortcomings (**R02**, **R07**, **R09**). Criticism focused especially on *Body* and *Face/Head*, while many other suggestions were mentioned only once, indicating individual preferences across players. Building on Harper's [16] argument that body-related customization concerns which bodies can be represented, our results specify this issue for a contemporary RPG editor: participants criticized limited *Body*-related customization not only because they wanted more *Body Type* options, but also because they wanted finer control over body proportions (**R02**, **R07**, Table 4). This points toward the need for broader representational coverage, more diverse options, and greater degrees of control. Beyond aesthetic variation, such limitations can also constrain identity expression and the character concepts players are able to create [25, 42]. Based on self-determination theory (SDT) [37] and research on meaningful choices (e.g., [31]), offering more options may support autonomy and intrinsic motivation, but also requires interfaces that keep browsing manageable. Other feature suggestions were rarer (**R03**), but the theme *Removing Restrictions* also relates to autonomy, as some participants want to remove game-mechanic or aesthetic restrictions (Table 1). While one can imagine removing these (e.g., making *Tails* available for every character), the game-mechanic-related suggestions would contradict the D&D rules (e.g., selecting spells from every class). However, in analog games, like tabletop role-playing, players can decide to alter rules within their game (so called *House rules* [22]). Mod usage can be seen similarly (**R02**, **R07**, **R09**).

6.2 Category Importance, Option Sufficiency, and Game Context

Regarding **RQ 2**, nearly all categories with direct game impact were assessed as important (**R05**), and several highly rated categories were aesthetic rather than mechanical. While we expected differences between aesthetic-only and game-impacting categories, the EFA revealed that the aesthetic categories can be differentiated further (**R10**). Notably, *Hairstyle*, *Body Type* and *Hair Color* have no direct game impact and belong to the same aesthetic factor, and were still among the most important categories (**R05**, **R10**). Building on McArthur's [28] observation that character customization is shaped by the concrete creation context, our results show that perceived category importance in a long RPG cannot be explained by game impact alone. Participants also valued categories because they supported imagined characters, role-playing, diversity, creative freedom, self-representation, and aesthetic relevance (**R06**; Table 2). We hypothesize that character-defining aesthetic elements become important if they are prominently visible even in the third-person view²¹, recognizable, and tied to the player's imagined character throughout the game. Consistent with prior work showing that customization can support self-representation, character attachment, identity exploration, and narrative intent [3, 15, 49], our results show how these broader motivations translate into the perceived importance of concrete customization categories.

The number of selection options in many categories is not sufficient (**R09**). However, some categories may require not only more presets, but also finer-grained controls, especially for face and body customization. Ducheneaut et al. [5] already showed in 2009 that *Hairstyle* and *Hair Color* were highly relevant and had insufficiently options in earlier virtual worlds. Our results replicate this pattern in a contemporary, more fine-grained RPG editor, but add that it persists despite substantially larger option sets in these categories (**R05**, **R09**, Table 5). This suggests that increasing option counts alone does not necessarily resolve perceived insufficiency. These findings

²¹Showing the character from an outside perspective, with the camera positioned behind or elsewhere in the surroundings.

add to the assumed tension (see introduction) between expressive depth and choice overload. Our data suggests that further increasing the options is possible, as participants rarely rated existing categories (even large ones) as offering too many options (R08). Yet, it raises the question of whether simply adding more options can satisfy players' expectations. Our data cannot show whether a point of full satisfaction exists, or whether highly invested players will continue to want more options as long as customization remains meaningful to them. As our study focused on a single game, the exact ranking of categories cannot simply be generalized to other character editors, as the context might be different [28]. Some findings are likely shaped by the specific RPG context of *BG3*, especially the importance of game-impacting categories such as *Class*, *Race*, and *Ability Points*. However, Ducheneaut et al. [5] investigated different online and virtual-world contexts, including *Maple Story*, *WoW*, and *Second Life*. The overlap with our findings therefore does not show direct generalizability, but suggests that some patterns, especially the importance and perceived insufficiency of hair-related customization may not be specific to *BG3*.

6.3 Individual Differences

Regarding RQ 3, the findings show that category importance varies across players and player groups (R11–R16). Even categories with low average importance should therefore not be dismissed, as they may still matter to specific groups. Previous studies have primarily focused on how individual differences relate to created characters or the selection of customization options [5, 6, 11, 12]. We extend this body of knowledge, as our study showed that individual differences become visible already at the level of perceived category importance, even before players select concrete options or produce a final character.

Non-binary and female participants reported more difficulty creating intended characters (R11, R12). This aligns with Løfgren Elvig and Mekler's [8] observation that character editors can constrain non-binary gender expression, and with Björnsson et al.'s [3] finding that the importance of character gender and representation differs across player groups. Given that these groups rate aesthetic details as more important than male participants (R13), we can hypothesize that not having enough options within these might be more critical for them. Age, personality, and engagement with the game further suggest that customization needs differ across player groups (R14–R16). While prior work linked individual differences to character creation and self-presentation [5, 6, 12], our findings suggest that such differences may already shape how players evaluate customization categories. However, some subgroup comparisons were limited by skewed distributions, so these findings should be interpreted as exploratory rather than definitive. Taken together, these differences (R11–R16) suggest that character editors should not be treated as "one size fits all" systems. Rather, future research should investigate individualized or adaptive editors that account for player differences and make personally relevant categories easier to access without forcing every player through the same customization path.

6.4 Design Suggestions

Based on our findings, we derive suggestions for future character editors (see Table 7). These suggestions are most directly grounded in the context of *BG3* and are therefore particularly relevant for similar, single character-centered RPG editors, where the created character remains central to gameplay, role-playing, and narrative engagement. For other editor types or genres, they should be understood as transferable starting points rather than universal design rules, as category importance and option sufficiency depend on the concrete game and editor context.

Table 7. Design suggestions for future character editors derived from the findings of this paper.

Focus	Empirical basis	Suggestion
Category prioritization	Some categories were more important than others (R05). <i>Body Type</i> and <i>Hairstyle</i> were also among the categories most often perceived as offering too few options (R09; Table 5).	Prioritize highly important rated categories when development resources are scarce and extend them through more diverse options, more depth, or finer-grained controls.
Optional customization depth	Participants were not generally overwhelmed by large option sets (R08), but individual differences show that players do not value all categories equally (R11–R16).	Offer deeper customization without making it mandatory to use, for example through presets, optional advanced controls, search and filtering, clearer category organization, and direct element selection.
Body and face representation	Participants repeatedly criticized limited <i>Body</i> and <i>Face/Head</i> customization (R02, R07). This also aligns with Harper's [16] argument that body customization concerns which bodies can be represented.	Provide more inclusive, diverse and fine-grained body and face controls, such as body proportions, height, weight, muscle amount, and face/head sliders or comparable adjustment systems.
Reduction of restrictions	Participants suggested unlocking existing options such as <i>Tails</i> , <i>Horns</i> , or color options more broadly (R03; Table 1).	Avoid restrictions. If these impact mechanics or narrative consistency users might be warned before changing. Broader access to existing assets can increase autonomy.
Support for expressiveness, individuality and role-playing	Participants valued customization categories because they supported <i>Creating Imagined Character</i> , <i>Direct Game Impact</i> , and <i>Aesthetic Relevancy</i> (R06; Table 2). Participants and mod usage also pointed to missing categories and features (R02, R03, R07, R09).	Add categories and features that support distinct character concepts, such as body tattoos, starting gear, gait, posture, wings, character portrait poses, camp clothes in the editor, and hair length adjustment.
Integration with game context	Participants valued categories not only for aesthetic reasons, but also for their <i>Direct Game Impact</i> (R06; Table 2). At the same time, highly important aesthetic categories such as <i>Hairstyle</i> , <i>Body Type</i> , and <i>Hair Color</i> show that not every category needs direct game impact (R05, R10).	Increase the relevance of categories either through clearly noticeable game impact (e.g., dialogue options, NPC reactions, or narrative consequences), or through aesthetic prominence. Game impact should be graspable, but does not need to be added to every category.

7 Limitations

We did not include all available customization categories of *BG3* to keep the survey duration length acceptable and not overwhelm participants. Across the free text answers, we learned that *Face/Head* was often criticized, so we can only assume an importance similar to *Body*, which was also mentioned frequently in the free text answers. This limitation needs to be kept in mind when arguing about the most important categories in *BG3*. Our sample—likely because of the popularity of the game and how we advertised our survey—consisted of only a few participants who had little experience with the game (only eleven had a playing time below 51h). Thus, results that we cover are mainly related to players that are invested in the game and are potentially more interested in the character editor, and there is a risk of survivorship bias: those who have not invested much in the game/editor might have stopped playing it because of issues with, for example, the editor itself. The perspective of those players is underrepresented in our data. We might also not have attracted players that are in general not invested much in character creation. Larian Studios provided a statistic in 2023, stating that 94% of their player base created a custom character (instead of using predefined/random ones) and that players have spent combined more than 8000 years in the editor²², so it is not clear how large this group actually is. The survey attracted more younger participants and the sample had mainly a Western nationality. These aspects limit generalizability and require future work focusing on the underrepresented participants and comparisons with our results. Another aspect is the distribution of the Big Five traits in our sample. Although the internal distribution was skewed, limiting our analysis to neuroticism, the overall trait levels did not substantially deviate from previously reported distributions in North American and Western European samples [38]. Minor differences, such as slightly higher neuroticism and somewhat lower extraversion scores, are consistent with more recent findings following the COVID-19 pandemic [48]. We did not collect

²²<https://x.com/larianstudios/status/1732091568243229159/photo/1>, last accessed: 07.07.2026

participants' experience with tabletop RPGs in general nor with D&D in particular. Potentially, experienced players might have a different view on character creation. Methodologically, it needs to be kept in mind that the character editor assessment questions were presented after participants assessed the 29 customization categories. This could have had an impact. We only asked for reasons why participants find their top three categories important, and not every category they assessed above the scale's mid-point. This was done to keep an already long survey doable for participants. Similarly, we did not ask for reasons why they found something unimportant.

8 Conclusion & Future Work

We analyzed how players perceive the *BG3* character editor and how they value offered customization categories. We found that although the character editor is well perceived overall, it has potential to improve. Criticism was mainly related to certain too-restricted customization categories and the wish for more options within categories, and less so for actual missing features. Across our results, we decided to not only present frequently raised aspects, but also category and feature suggestions that originated from only a low number of participants, as a resource for future character editor development. This paper provides an overview on which categories participants value in a fantasy video game context, and also provides design suggestions, and can thus serve as guide for character editors in similar contexts. We also learned that it is apparently not harmful to offer hundreds of options and given our results, it is reasonable to do so from a player's perspective. We were able to find perception differences already on the level of customization categories depending on age, gender, personality traits and engagement with the game, adding to the body of knowledge in this research area.

From here, many paths for future work open up: First, re-doing our study within other fantasy RPGs and/or in non-fantasy contexts would be interesting to compare and complement results to learn about generalizability (as we replicated parts of Ducheneaut et al.'s [5] work, we expect general patterns across genres, game types and character editor designs). Some games with, from the participants' perspective, better editors were also mentioned multiple times and should be focused on first. This would also be necessary to derive verified overarching design guidelines for character editors, which are currently unavailable to our knowledge. For this, our design suggestions can be a starting point. It would be interesting to also investigate differences between editors of games played mainly in first²³ vs. third person view, and to consider editors that offer more continuous customization categories (e.g., face morphing sliders). Although participants are apparently not overwhelmed with categories with hundreds of elements, the question arises what happens if many categories have so many selection options. This raises questions towards choice overload [20] and it requires sophisticated user interface features. Future work can investigate which features would specifically be necessary in such cases. Our study focused purely on the players' viewpoint. That *BG3* and other editors impose certain restrictions might be necessary from the developers' perspective. Future work should investigate this, for example, through a series of interviews. It is likely that especially in comparisons between AAA and Indie productions [24], many aspects that would be reasonable from a players' perspective are impossible to realize with the budgets companies have today. But even here, player-centric work like ours can help to prioritize well. For investigating perception differences (RQ 3), we restricted ourselves to four discriminating factors so as not to overwhelm participants with somewhat unusual personal questions (given the scope of the survey, which also accounted for two other research questions) and where we assumed that these would not lead to irritation as these are commonly asked for in surveys. Future work, in focused studies, could consider other factors, for example, skin color (based on [5]), or

²³View from the character's own perspective at eye height.

sexuality [3]. Investigating which mods are offered for what and which mods are frequently used by participants would be an interesting next step, that could complement the view on important customization categories. Another future work opportunity is to investigate how character creation tools affect the later player experience, which we did not focus on in this paper. Our data shows which categories participants considered important during creation, but it does not show how these choices specifically shape subsequent play, identification with the character, narrative engagement, or long-term satisfaction with the creation. Two future work directions were mentioned previously: First, it would be interesting to do a similar study primarily targeting players that decided to stop playing *BG3* early on, to learn about their reasons (and especially, whether these are character editor-related), and to compare their importance ratings to those of this survey. Second, one could focus on tailored character editors, i.e., editors that consider the individual differences found and for example re-arrange customization categories according to them.

References

- [1] Konstantinos C. Apostolakis and Petros Daras. 2013. RAAT – The Reverie Avatar Authoring Tool. In *2013 18th International Conference on Digital Signal Processing*. IEEE, 1–6. <https://doi.org/10.1109/ICDSP.2013.6622788>
- [2] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B* 57, 1 (1995), 289–300. <https://www.jstor.org/stable/2346101>
- [3] Gunnar H. Björnsson, Maximilian J. Dechant, and Susanne Poeller. 2026. Queering Character Creation: Player Perspectives on Choosing Characters' Gender and Sexual Orientation in Role-Playing Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1–17. <https://doi.org/10.1145/3772318.3791047>
- [4] Johan Braeken and Marcel A. L. M. Van Assen. 2017. An Empirical Kaiser Criterion. *Psychological Methods* 22, 3 (2017), 450–466. <https://doi.org/10.1037/met0000074>
- [5] Nicolas Ducheneaut, Ming-Hui Wen, Nicholas Yee, and Greg Wadley. 2009. Body and Mind: A Study of Avatar Personalization in Three Virtual Worlds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1151–1160. <https://doi.org/10.1145/1518701.1518877>
- [6] Robert A. Dunn and Rosanna E. Guadagno. 2012. My Avatar and Me: Gender and Personality Predictors of Avatar–Self Discrepancy. *Computers in Human Behavior* 28, 1 (2012), 97–106. <https://doi.org/10.1016/j.chb.2011.08.015>
- [7] Rob Eisinga, Manfred te Grotenhuis, and Ben Pelzer. 2013. The Reliability of a Two-Item Scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health* 58, 4 (2013), 637–642. <https://doi.org/10.1007/s00038-012-0416-3>
- [8] Markus Løfgren Elvig and Elisa D. Mekler. 2026. Often Discouraged, Never Refused? How Character Creator Interfaces Afford Non-Binary Gender Expression. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1–17. <https://doi.org/10.1145/3772318.3790841>
- [9] Leandre R. Fabrigar and Duane T. Wegener. 2012. *Exploratory Factor Analysis*. Oxford University Press. <https://doi.org/10.1093/acprof:osobl/9780199734177.001.0001>
- [10] Andy Field, Zoe Field, and Jeremy Miles. 2012. *Discovering Statistics Using R*. SAGE. <https://collegepublishing.sagepub.com/products/discovering-statistics-using-r-1-236067>
- [11] Emmanuel Fokides. 2021. My Avatar and I. A Study on Avatars, Personality Traits, Self-Attributes, and Their Perceived Importance. *Journal of Ambient Intelligence and Humanized Computing* 12, 1 (2021), 359–373. <https://doi.org/10.1007/s12652-020-01977-1>
- [12] Katrina Fong and Raymond A. Mar. 2015. What Does My Avatar Say About Me? Inferring Personality From Avatars. *Personality and Social Psychology Bulletin* 41, 2 (2015), 237–249. <https://doi.org/10.1177/0146167214562761>
- [13] Lewis R. Goldberg. 1999. A Broad-Bandwidth, Public Domain, Personality Inventory Measuring the Lower-Level Facets of Several Five-Factor Models. *Personality Psychology in Europe* 7, 1 (1999), 7–28. <https://admin.unt.edu.pk/Media/Site/STD/FileManager/OsamaArticle/26august2015/A%20broad-bandwidth%20inventory.pdf>
- [14] Mary E. Mc Grath. 2025. *Queernormativity & Freedom in Baldur's Gate 3: Enabling True Self Exploration & Expression for Queer Players*. Ph.D. Dissertation. Institute of Art, Design + Technology. <https://hdl.handle.net/10779/iadt.30610784>
- [15] Yisong Han and Xavier Ho. 2024. “More Than Just Playing a Character”: Gender Exploration and Expression in Avatar Customization. *Games and Culture* 21, 5 (2024), 1–23. [doi:10.1177/15554120241307982](https://doi.org/10.1177/15554120241307982)
- [16] Todd Harper. 2020. Endowed by Their Creator: Digital Games, Avatar Creation, and Fat Bodies. *Fat Studies* 9, 3 (2020), 259–280. <https://doi.org/10.1080/21604851.2019.1647022>

- [17] Amy A. Hasinoff and Rena Bivens. 2021. Feature Analysis: A Method for Analyzing the Role of Ideology in App Design. *Journal of Digital Social Research* 3, 2 (2021), 89–113. <https://doi.org/10.33621/jdsr.v3i2.56>
- [18] John L. Horn. 1965. A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrika* 30, 2 (1965), 179–185. <https://doi.org/10.1007/BF02289447>
- [19] Hsiu-Fang Hsieh and Sarah E. Shannon. 2005. Three Approaches to Qualitative Content Analysis. *Qualitative Health Research* 15, 9 (2005), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- [20] Sheena S. Iyengar and Mark R. Lepper. 2000. When Choice Is Demotivating: Can One Desire Too Much of a Good Thing? *Journal of Personality and Social Psychology* 79, 6 (2000), 995–1006. <https://doi.org/10.1037/0022-3514.79.6.995>
- [21] Henry F. Kaiser and John Rice. 1974. Little Jiffy, Mark IV. *Educational and Psychological Measurement* 34, 1 (1974), 111–117. <https://doi.org/10.1177/001316447403400115>
- [22] Ville Kankainen and Janne Paavilainen. 2019. Hybrid Board Game Design Guidelines. In *Proceedings of DiGRA 2019 Conference: Game, Play and the Emerging Ludo-Mix*. 1–12. <https://doi.org/10.26503/dl.v2019i1.1098>
- [23] Ben Light, Jean Burgess, and Stefanie Duguay. 2018. The Walkthrough Method: An Approach to the Study of Apps. *New Media & Society* 20, 3 (2018), 881–900. <https://doi.org/10.1177/1461444816675438>
- [24] Nadav Lipkin. 2013. Examining Indie’s Independence: The Meaning of “Indie” Games, the Politics of Production, and Mainstream Cooptation. *Loading... The Journal of the Canadian Game Studies Association* 7, 11 (2013), 8–24. <https://loading.journals.publicknowledgeproject.org/loading/index.php/loading/article/view/122>
- [25] Cecilie Lolansen, Nikki D. Marshall, Courteney T. L. Fisher, and Tyrone L. Burleigh. 2025. A Decade of Avatars: A Systematic Review of Recreational Gaming as a Mechanism for Gender Identity Exploration and Affirmation. *International Journal of Transgender Health* (2025), 1–24. <https://doi.org/10.1080/26895269.2025.2595471>
- [26] Philipp Mayring. 2014. *Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution*. Self-Published, 79–87. <http://nbn-resolving.de/urn:nbn:de:0168-ssolar-395173>
- [27] Victoria McArthur. 2017. The UX of Avatar Customization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Denver, CO, USA, 5029–5033. <https://doi.org/10.1145/3025453.3026020>
- [28] Victoria McArthur. 2019. Making Mii: Studying the Effects of Methodological Approaches and Gaming Contexts on Avatar Customization. *Behaviour & Information Technology* 38, 3 (2019), 230–243. <https://doi.org/10.1080/0144929X.2018.1526969>
- [29] Victoria McArthur and Jennifer Jenson. 2014. E Is for Everyone? Best Practices for the Socially Inclusive Design of Avatar Creation Interfaces. In *Proceedings of the 2014 Conference on Interactive Entertainment*. ACM, 1–8. <https://doi.org/10.1145/2677758.2677783>
- [30] Adam W. Meade and S. Bartholomew Craig. 2012. Identifying Careless Responses in Survey Data. *Psychological Methods* 17, 3 (2012), 437–455. <https://doi.org/10.1037/a0028085>
- [31] Scott Nicholson. 2012. A User-Centered Theoretical Framework for Meaningful Gamification: A User-Centered Theoretical Framework for Meaningful Gamification. *Proceedings of the 8th International Conference on Games + Learning + Society* 8 (2012), 1–7. <https://scottnicholson.com/pubs/meaningfulframework.pdf>
- [32] Jum C. Nunnally and Ira H. Bernstein. 1994. *Psychometric Theory* (3 ed.). McGraw-Hill.
- [33] Tyler Pace. 2008. Can an Orc Catch a Cab in Stormwind? Cybertype Preference in the World of Warcraft Character Creation Interface. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems*. ACM, Florence, Italy, 2493–2502. <https://doi.org/10.1145/1358628.1358706>
- [34] Johannes Pfau. 2025. Progression Balancing × Baldur’s Gate 3: Insights, Terms and Tools for Multi-Dimensional Video Game Balance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1–12. <https://doi.org/10.1145/3706598.3713162>
- [35] Katherine M. Rahill and Marc M. Sebrechts. 2021. Effects of Avatar Player-Similarity and Player-Construction on Gaming Performance. *Computers in Human Behavior Reports* 4 (2021), 1–11. <https://doi.org/10.1016/j.chbr.2021.100131>
- [36] Beatrice Rammstedt and Oliver P. John. 2007. Measuring Personality in One Minute or Less: A 10-Item Short Version of the Big Five Inventory in English and German. *Journal of Research in Personality* 41, 1 (2007), 203–212. <https://doi.org/10.1016/j.jrp.2006.02.001>
- [37] Richard M. Ryan, Edward L. Deci, et al. 2002. Overview of Self-Determination Theory: An Organismic Dialectical Perspective. *Handbook of Self-determination Research* 2 (2002), 3–33. <https://www.elaborer.org/cours/A16/lectures/Ryan2004.pdf>
- [38] David P. Schmitt, Jüri Allik, Robert R. McCrae, and Verónica Benet-Martínez. 2007. The Geographic Distribution of Big Five Personality Traits: Patterns and Profiles of Human Self-Description Across 56 Nations. *Journal of Cross-Cultural Psychology* 38, 2 (2007), 173–212. <https://doi.org/10.1177/0022022106297299>
- [39] Patrick Schuck, Maximilian Altmeyer, Antonio Krüger, and Pascal Lessel. 2022. Viewer Types in Game Live Streams: Questionnaire Development and Validation. *User Modeling and User-Adapted Interaction* 32, 3 (2022), 417–467. <https://doi.org/10.1007/s11257-022-09328-9>

- [40] Iris Seidinger and Jens Grubert. 2016. 3D Character Customization for Non-Professional Users in Handheld Augmented Reality. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE, Merida, Mexico, 129–134. <https://doi.org/10.1109/ISMAR-Adjunct.2016.0059>
- [41] Sally Sheppard. 2025. *Aesthetic Choices and Player Motivation in Baldur's Gate 3*. Master's Thesis. University of Kent. <https://kar.kent.ac.uk/id/eprint/111974>
- [42] Xan Smith. 2025. Binary Barriers: Avatar Creation and Play for Nonbinary Video Game Players. *Games and Culture* (2025), 1–22. <https://doi.org/10.1177/15554120251322209>
- [43] Katta Spiel, Oliver L. Haimson, and Danielle Lottridge. 2019. How to Do Better With Gender on Surveys: A Guide for HCI Researchers. *Interactions* 26, 4 (2019), 62–65. <https://doi.org/10.1145/3338283>
- [44] Steve Stemler. 2000. An Overview of Content Analysis. *Practical Assessment, Research, and Evaluation* 7, 1 (2000), 1–10. [doi:10.7275/z6fm-2e34](https://doi.org/10.7275/z6fm-2e34)
- [45] James P. Stevens. 2001. *Applied Multivariate Statistics for the Social Sciences* (4 ed.). Psychology Press. <https://doi.org/10.4324/9781410604491>
- [46] Thomas Studley. 2025. Meta-Music and the Sonic Storytelling of Baldur's Gate 3. *Convergence – Special Issue: Sound and Music for Immersive Storytelling* (2025), 1–26. <https://doi.org/10.1177/13548565251352073>
- [47] Louis K. P. Suryapranata, Gede P. Kusuma, Yaya Heryadi, Bahtiar S. Abbas, Lukas, and Adang S. Ahmad. 2017. Personality Trait Prediction Based on Game Character Design Using a Machine Learning Approach. In *International Conference on Innovative and Creative Information Technology*. IEEE, 1–5. <https://doi.org/10.1109/INNOCIT.2017.8319139>
- [48] Angelina R. Sutin, Yannick Stephan, Martina Luchetti, Damaris Aschwanden, Ji Hyun Lee, Amanda A. Sesker, and Antonio Terracciano. 2022. Differential Personality Change Earlier and Later in the Coronavirus Pandemic in a Longitudinal Sample of Adults in the United States. *PLoS One* 17, 9 (2022), 1–16. <https://doi.org/10.1371/journal.pone.0274542>
- [49] Kim Szolin, Daria J. Kuss, Filip M. Nuyens, and Mark D. Griffiths. 2023. “I Am the Character, the Character Is Me”: A Thematic Analysis of the User-Avatar Relationship in Videogames. *Computers in Human Behavior* 143 (2023), 1–13. <https://doi.org/10.1016/j.chb.2023.107694>
- [50] Cathy Urquhart, Yvonne Underhill-Sem, Tyler Pace, Alain Houssian, and Victoria McArthur. 2009. Are Socially Exclusive Values Embedded in the Avatar Creation Interfaces of MMORPGs? *Journal of Information, Communication and Ethics in Society* 7, 2/3 (2009), 192–210. <https://doi.org/10.1108/14779960910955909>
- [51] Li-Fang Zhang. 2002. Thinking Styles and the Big Five Personality Traits. *Educational Psychology* 22, 1 (2002), 17–31. <https://doi.org/10.1080/01443410120101224>

Ludography

- [G1] *Baldur's Gate 3*. Larian Studios / Wizards of the Coast, 2023. Game [Windows, macOS, Linux, PlayStation 5, Xbox Series, GeForce Now].
- [G2] *Black Desert*. Pearl Abyss, 2014. Game [Windows, PlayStation 4, PlayStation 5, Xbox One, Xbox Series, GeForce Now].
- [G3] *Cyberpunk 2077*. CD Projekt, 2022. Game [Windows, macOS, PlayStation 4, PlayStation 5, Xbox One, Xbox Series X/S, Nintendo Switch 2, Stadia].
- [G4] *Dragon Age: The Veilguard*. Electronic Arts, 2024. Game [Windows, PlayStation 5, Xbox Series, GeForce Now, xCloud].
- [G5] *Elden Ring*. FromSoftware, 2022. Game [Windows, PlayStation 4, PlayStation 5, Xbox One, Xbox Series, Nintendo Switch 2].
- [G6] *Jam City Roller girls*. Frozen Codebase, 2011. Game [Wii].
- [G7] *Maple Story*. Nexon, 2003. Game [Windows, Android, GeForce Now].
- [G8] *Mii Maker*. Nintendo, 2006. Game [Nintendo 3DS, Wii, Wii U, Nintendo Switch].
- [G9] *Neverwinter Nights 2*. Obsidian Entertainment and Atari SA, 2000. Game [Windows, macOS, PlayStation 5, Xbox Series, Nintendo Switch].
- [G10] *Saints Row 2*. THQ, 2008. Game [Windows, PlayStation 3, Xbox 360, Linux, GeForce Now, Amazon Luna].
- [G11] *Second Life*. Linden Lab, 2003. Game [Windows, macOS, Linux].
- [G12] *The Elder Scrolls V: Skyrim*. Bethesda Software, 2011. Game [Windows, PlayStation 3, PlayStation 4, PlayStation 5, Xbox 360, Xbox One, Xbox Series X/S, Nintendo Switch, Nintendo Switch 2].
- [G13] *The Sims 4*. Electronic Arts, 2014. Game [Windows, macOS, PlayStation 4, Xbox One].
- [G14] *World of Warcraft*. Blizzard Entertainment, 2004. Game [Windows, macOS].

A Appendix

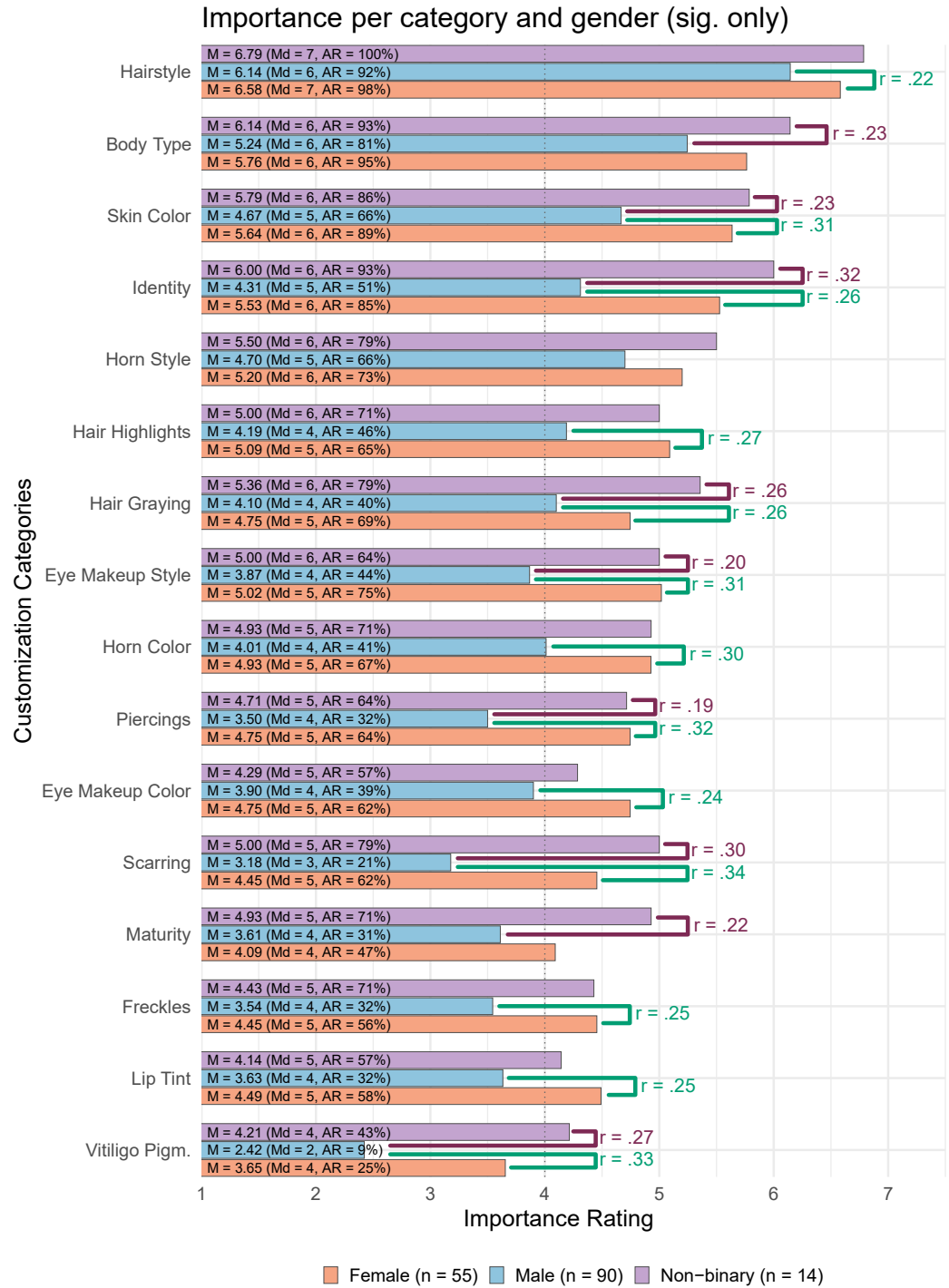


Fig. A1. Comparison of customization categories' importance ratings between genders. Shows only categories with significant differences. Mean (M), median (Md), effect size (r), agreement rate (AR), and significant post-hoc tests differences (connecting lines between groups).

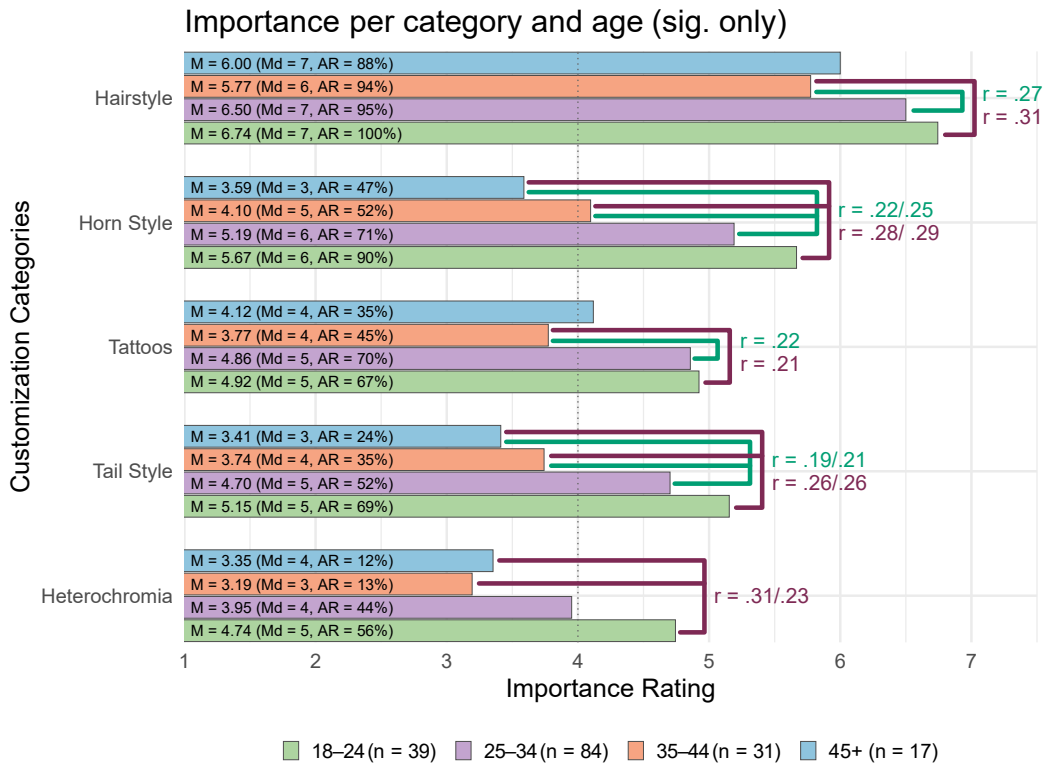
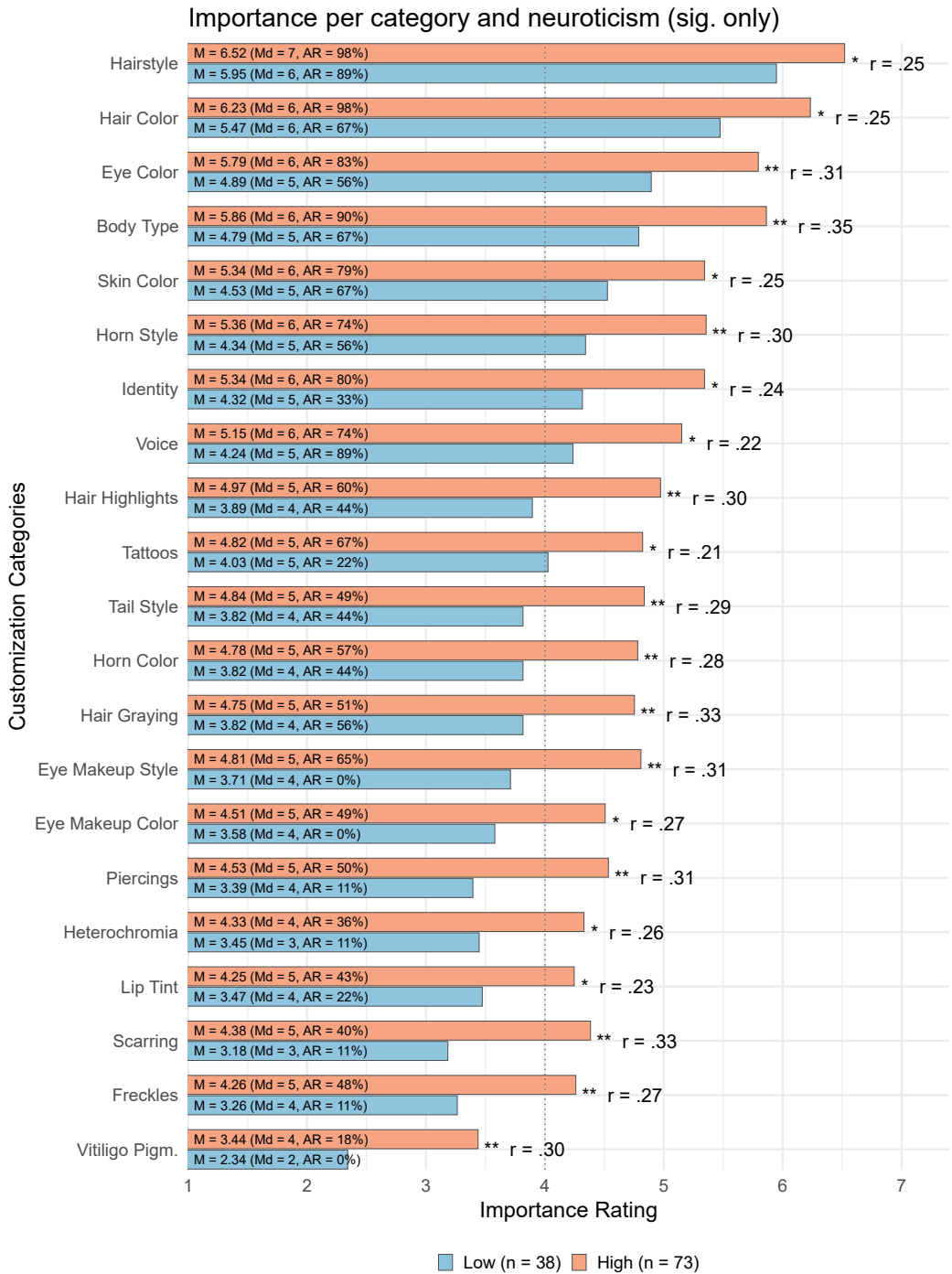


Fig. A2. Comparison of customization categories' importance ratings between age groups. Shows only categories with significant differences. Mean (M), median (Md), effect size (r), agreement rate (AR), and significant post-hoc tests differences (connecting lines between groups).



** p < .01, * p < .05

Fig. A3. Comparison of customization categories' importance ratings between low and high neuroticism. Shows only categories with significant differences. Mean (M), median (Md), effect size (r), agreement rate (AR), and significant difference (*).

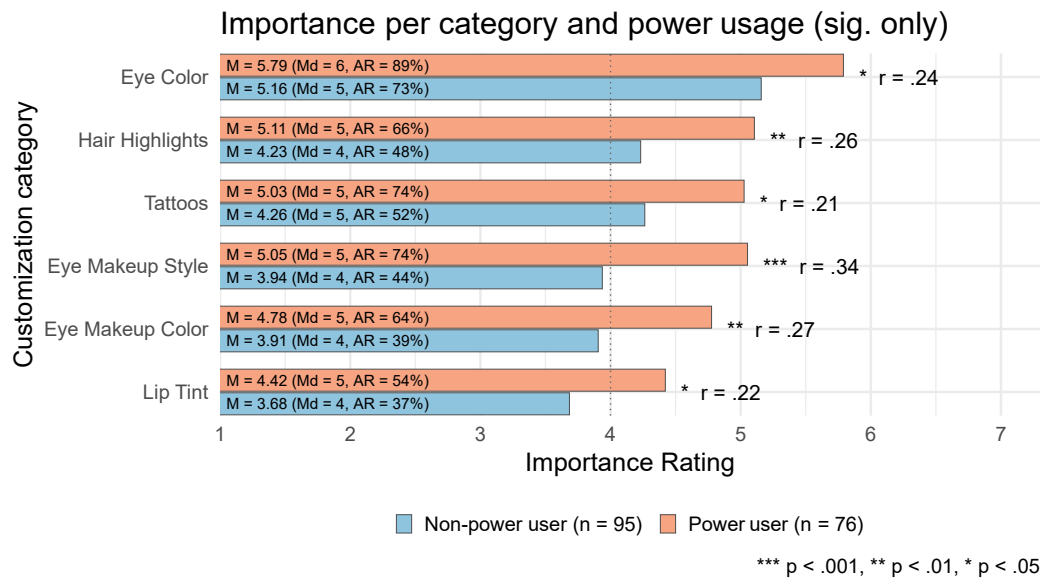


Fig. A4. Comparison of the customization categories' importance ratings between power users and non-power users. Shows only categories with significant differences. Mean (M), median (Md), effect size (r), agreement rate (AR), and significant difference.

Table A1. Reasons (why a category is important) × customization categories. Categories without free text answers: *Horn Color, Hair Graying, Tail Style, Scarring, Piercings, Eye Makeup Color, Lip Tint, Hair Highlights, Vitiligo Pigmentation*. **N** = Number of participants who were asked the respective question; **n** = Number of participants who responded. **X** = Total number of codes assigned.

Code n/N	X	Creating Imagined Character																	
		Class 103/122	Race 58/72	HairStyle 64/71	Ability Points 50/60	Hair Color 17/22	Body Type 19/25	Eye Color 12/13	Identity 11/16	Background 9/10	Voice 8/8	Skin Color 9/6	Tattoos 6/6	Facial Hairstyle 6/8	Horn Style 2/3	Maturity 2/3	Freckles 2/2	Genitals 2/2	Vitiligo Pigmentation 0/1
Relevancy for character																			
90	16	5	28	14	4	5	2	1	3	5	3	1	1	1	1	0	0	0	
Role-playing																			
53	19	17	1	5	0	0	1	1	6	1	0	1	0	0	0	0	1	0	
Unique touches																			
39	3	3	12	5	4	1	3	0	0	2	0	3	3	0	0	0	0	0	
Diversity																			
24	0	0	1	0	0	9	0	7	0	0	4	0	0	0	2	0	1	0	
Fiction																			
15	0	11	2	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	
Creative freedom																			
10	1	2	1	4	0	1	0	0	0	0	0	0	0	1	0	0	0	0	
Recreate self																			
4	0	0	3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
Direct Game Impact																			
Gameplay																			
138	67	32	0	36	0	0	0	1	2	0	0	0	0	0	0	0	0	0	
Replayability																			
27	18	5	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Appearance less relevant																			
19	9	4	0	5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	
Adding immersion																			
9	0	0	1	0	1	2	0	1	0	2	2	0	0	0	0	0	0	0	
Aesthetic Relevancy																			
Aesthetic reasons																			
46	0	8	20	0	3	4	3	0	0	1	4	2	1	0	0	0	0	0	
Visibility																			
30	0	3	15	0	6	0	3	0	0	0	1	1	1	0	0	0	0	0	
Relevancy for other elements																			
23	2	4	4	6	0	2	0	0	0	2	1	1	1	0	0	0	0	0	
Exposure time																			
9	0	0	4	0	3	0	0	0	0	1	1	0	0	0	0	0	0	0	

Received February 2026; revised June 2026; accepted July 2026