AI Musician, Do You Have a Place in this World? An Empirical Study

KEVIN LAKSANA ISKANDAR, University of Twente, The Netherlands

Abstract

The use of Artificial Intelligence (AI) technology has mostly been applied to several processes and areas of music namely; composition, improvisation, performance, education & listening. However, their application is not always accepted by humans. This study focuses on the composition and performance aspect of AI musicians and investigates how human perception on AI musicians may impact one's evaluation of AI music. This study aims to investigate how the perception of individuals towards artificial intelligence may influence the evaluation of AI performers and AI musicians. By using the grounded theory approach, a literature review was performed and as a result several concepts pertaining to AI musicians were observed. These concepts being; anthropomorphism, capabilities, fear, prejudice, bias and acceptability. A scale measuring human perception of AI musicians and performers (PREAMP) was developed based on the study. Furthermore, a qualitative study was performed generating a total of 122 respondents to investigate the relationship even further. The results suggest the correlation of several constructs namely; AI acceptance & capability, perception of AI musician & evaluation of AI music, and capability & evaluation of AI music. The results suggest a positive correlation between the evaluation of AI music and one's perception of AI musicians & performers, and that the overall quality of AI music is lower than that of human music.

Additional Key Words and Phrases: Music, AI, Artificial Intelligence, Bias, Perception

1 INTRODUCTION

"AI is undoubtedly superior to humans in identifying new combinations of attributes, and hence can create more novelties, but this does not mean that the novelty will be socially desirable." [30] One such novelty that an AI may create is music. Music is constantly evolving. New genres of music are formed and the popularity of different genres shift over time. In the late 1900s, the genre of Rock and RBSoul was at the pinnacle of popularity but now Rap/Hip-hop and Pop has taken its place. [5] There are several theories surrounding the evolution of music. It may be a result of sexual selection, influence of social groups and caregiving [10]. However, one reason that stands out is *"the importance of new recording technology on genre preference"* [5]. This would suggest that the introduction of new technological advancements may change the trend of music. One of those technological advancements applied to the field of music is the use of Artificial Intelligence.

The use of Artificial Intelligence (AI) technology has mostly been applied to several processes and areas of music namely; composition, improvisation, performance, education & listening.[9, 12] In the area of composition, AI is commonly used to generate music based on the provided dataset of existing music. Machine learning techniques such as deep learning and neural networks are often used to generate melodies, harmonies and music structure which results in the composition of music.[2] AI may also be used to generate musical improvisation on an existing piece of music based on style and input provided. [12] Aside from that, AI is also used for music education, be it in the context of academics or personal use for music practice/tutoring, there exists several renowned programs and applications for this purpose.[9] There also exists several applications which may identify music by listening to a small portion of audio. In regards to performance, AI may generate audible output based on the provided input. This may range from simple audible output of notes from a Musical Instrument Digital Interface (MIDI) instrument [12] to the performance of advanced 3-Dimensional choreography and vocals of a whole musical piece.[3,17] AI music performance systems may also be used as a supporting tool and an enabler used by musicians and performers to improve or enhance the musical performances such as the use of holograms to accompany performers and the ability to control a piano based on a dancer's movement allowing the music to adjust to a dancer's movements during a performance instead of the contrary. [9,19]

Some are skeptical towards the use of AI in music claiming that a machine is not fit for the task[9] as they are just producing a copy of what humans do and are not to be considered genuine music. However, what if there exists a sentient AI which possesses a consciousness similar to that of a human with its ability to perform all the functions of composition and performance. Will this AI be able to perform to an audience just like human performers do and produce genuine music? This however is extremely difficult to answer as the appraisal of both art and AI consciousness is very difficult due to a lack of clear definition. [11] Thus one cannot help but wonder whether AI may one day create a new genre in music, can it replace humans as performers and composers? This evolution might already be happening with the introduction of Virtual/AI performers and the use of AI composers. "When compared to the human ability for music composition, algorithms' capacity to generate outputs is faster, more scalable and cheaper."[4] The use of AI may bear many benefits both known and yet to be known, but how does human society perceive this change? May it induce bias towards the use of AI as it is believed that machines may not make music like humans do, may the use of AI be supported

or may some other result be the outcome? With all this in mind, the question is posed; do you have a place in this world?

Numerous different research has been done in the field pertaining to the use of AI in the domain of music. Although there may not be a shortage of research in this area, many focus on the technical spectrum of the field and there is not much in the area of a human's perception or in this case, that of a music listener. [4] Since 2020, several research has been done regarding a human's perception on AI-composed music [13, 14, 15, 16] and although results of different studies may vary, nevertheless, it has been explored in several instances. However, much a focus is given specifically towards the domain of AI-composed music but there is a lack of research pertaining to the perception of human society towards AI music performance systems. Thus, this paper aims at bridging the literature gap by also investigating human perception on AI music performance.

For the investigation to take place, this study will firstly conduct a systematic literature review using the grounded theory approach [22] to gather information and related studies relevant to the subject in order to build an improved understanding and viewpoint of the subject. Based on the findings of the literature review, the constructs that may influence the perception of humans towards AI musicians and performers will be identified and used to form a scale to measure an individual's perception of AI musicians and performers. A quantitative study will then be conducted to validate the scale and test the relationships suggested by the literature study. (*See Appendix D.7*)

1.1 Research Question

Main research question:

How does the perception of humans/music listeners towards artificial intelligence influence the evaluation of AI performers and AI musicians?

2 LITERATURE REVIEW

A literature review will be conducted via Scopus using the initial search string "music AND (ai OR (artificial AND intelligence)) AND (bias OR perception)". The use of grounded theory [22] will be applied into the literature review process to ensure its clarity, reproducibility and replicability.

2.1 Define

Several inclusion and exclusion criterias are defined. As AI is a field that advances as technology advances and new and improved methods of machine learning are introduced having an impact on AI-music technology, the inclusion criteria of 2020 or above will be set. This is also due to a previous observation in which papers relevant to the study are published from 2020 onwards. Although there may be interesting papers published in earlier years such as 2002 [12], the systems and technology are now outdated and society's perception of music and AI may change overtime[5] causing older literature to be less relevant. The inclusion criteria will also include only the use of journal articles or conference papers as they are a reliable source of information due to cross-checking and the lack of bias reviews. The papers must also be in English. The study will exclude several subject areas namely; Physics and Astronomy, Medicine, Health Professions, Chemical Engineering, Biochemistry, Genetics and Molecular Biology, Nursing and Neuroscience as these fields are outside the scope of this study.

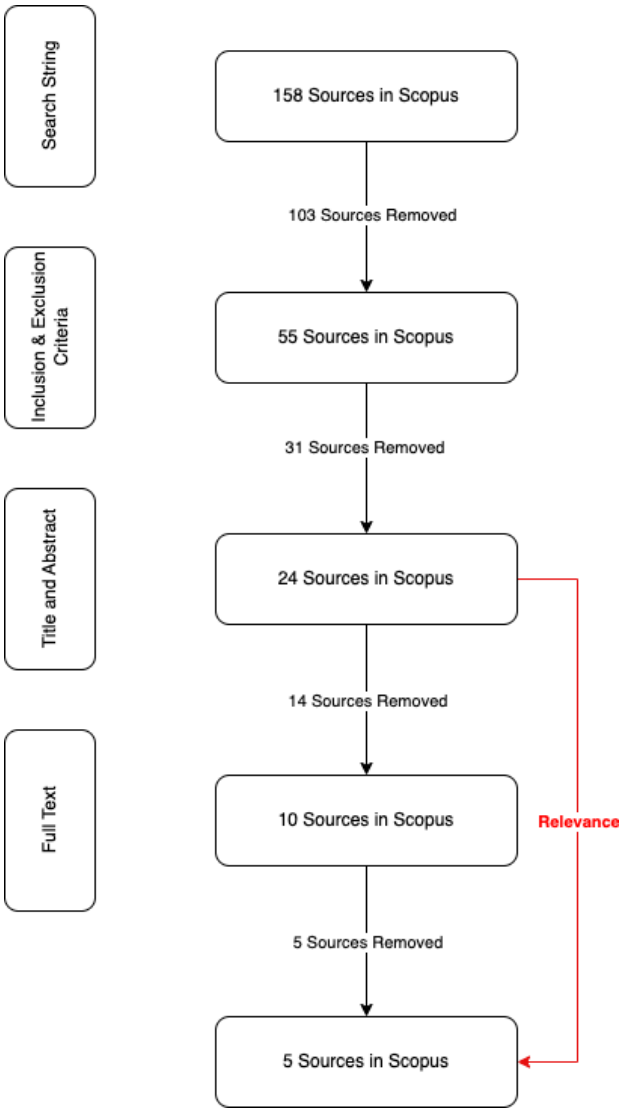


Figure 1. PRISMA Model of Search Process

2.2 Search

Upon initial search on scopus using the search string, removal of duplicates, and application of inclusion and exclusion criteria resulted in 79 papers. However, it was soon observed that the search string used did not cover the whole scope of the study. Some terms such as perception were also too general giving out irrelevant results. Numerous irrelevant sources were present and the area of AI use in music performance was not present. Thus a new search string was used in order to include all relevant topics. The search string being: "(music OR ((music OR musical) AND performance)) AND (ai OR (artificial AND intelligence) OR ((ai OR artificial) AND performer)) AND (bias OR ((human OR society) AND (perception)))". The study will now also add an inclusion criteria in which the subject area shall only include Arts & Humanities, Computer Science, Mathematics, Psychology, Social Sciences and Engineering as the new search string creates an abundance of unrelated subject areas. Computer Science and Engineering were included in the inclusion criteria although they presented papers focused on the technical aspects of the use of AI in music as some may be relevant.

2.3 Select

The search using the search string, and inclusion & exclusion criteria resulted in a total of 55 sources with no duplicates. Afterwards, in order to narrow down the number of sources, an observation of their relevance was made based on the title and abstract of the sources resulting in the elimination of 31 sources. From the 24 remaining sources, sources which were not highly relevant to the study were eliminated resulting in 10 remaining sources [24, 25, 26, 27, 28, 29, 30, 31, 32, 33] which served as candidates for the final sample of the systematic literature review. After a full text read of the 10, another 5 sources were removed resulting in a final sample of 5 papers. [25, 28, 29, 30, 33]

Although other sources discovered prior to the literature review were extremely relevant to the study, these sources were not included as they were not part of the results of the search string, resulting in a relatively small sample size of 5 sources. Numerous relevant sources did not include the keyword "perception" or "bias" but only included "Artificial Intelligence" and "music" causing the sources to be filtered out during the *search* phase. The removal of the keywords "perception" or "bias" from the search string resulted in 1024 sources for the title and abstract review stage after the inclusion and exclusion criteria have been applied with most of them being irrelevant to the study. Thus, due to the excessive amount of sources if the search string was made to be less specific, the study will proceed with the 5 sources.

The questions regarding the use of Artificial Intelligence does not only span in the domain of music, it is also observable in other topics of discussion such as art, healthcare, human resource management (HRM), human rights and intelligence in general. The concerns pertaining to the use of AI in the domain of art are similar to that in the domain of music in the sense that some may be skeptical towards the idea of AI performing a "human task". [39] Concerns may differ in the healthcare industry, where there is a lack of trust in AI systems. This is due to the possibility of overfitting & bias in AI decision-making and the high risk, of at worst, one's life. [37] Thus, patients may prefer human physicians over AI. Diversely, in the area of HRM there are concerns regarding the accountability of fairness of AI systems along with ethical and legal implications of AI-supported decision making by management. [41] This may also contribute to the significant impact AI has on human rights, in which the effects are more negative than positive in various rights, namely: privacy, equality, free expression, association, assembly, and work. [40] The trustworthiness, ethics and explainability [38] of Artificial Intelligence is in question. Furthermore, the novelty and ingenuity of AI creations are doubted. Although there may be exceptions, people seem to be skeptical of AI creativity in general.

2.4 Concepts

Sources/Themes	Anthropomorphism	AI Musician Capabilities	Prejudice/Stereotype Towards AI Musicians	Fear Towards AI	Biased Towards Humans	Relation Between AI Acceptability and AI Music Evaluation
Hong et al. (2022) [25]	✓	✓	✓	-	✓	✓
Sajadieh (2023) [28]	✓	✓	-	✓	-	✓
Sun et al. (2023) [29]	✓	✓	✓	-	✓	✓
Tubadji et al. (2021) [30]	-	✓	✓	✓	✓	✓
Déguernel et al. (2022) [33]	-	✓	✓	-	-	✓

Figure 2. Concept Matrix

The common concepts observed from the sources are shown in *Figure 2*. Although not all concepts are observed across all sources, a total of six most observable recurring concepts are selected. These concepts being *Anthropomorphism, AI Musician Capabilities, Prejudice/Stereotype Towards AI Musicians, Fear Towards AI, Biased Towards Humans*, and the *Relation Between AI Acceptability and AI Music Evaluatio*n. The scope and definition of each concept will be further elaborated in this section.

2.4.1 Anthropomorphism

Anthropomorphism is the practice of enabling non-human entities, such as robots or animals, to exhibit human-like form, characteristics, or behavior. [29] Thus as the definition suggests, this concept refers to sources mentioning the possibility of AI musicians having a human-like image or representation and its impact on human's perception on AI musicians. This image/representation may come in several different forms and has multiple levels of realism/human-like features. There are different degrees of anthropomorphism which may influence perception/acceptance. Low degree

using image/cartoon based representation for AI musicians. High degree using hyper-realistic human representations such as robots with human-like features or extremely human-like projections/representations. [29]

Anthropomorphism may cause humans to have higher expectations for AI music thus may be more harsh in evaluating AI music with higher degree of anthropomorphism. This is due to the fact that a higher degree of anthropomorphism causes humans to think that the AI musician is more competent. [29] The application of Anthropomorphism on AI musicians may increase acceptance of AI musicians if done right. This may be due to effectance and social motivation. Doing so may help make sense and better understand unfamiliar agents and reduce uncertainty. It also helps create a social connection. “Having anthropomorphic aspects, including being embodied and having human capacities, leads to the acceptance of its musician role. “ [25]

Making AI musicians look more lively and cute may help make people accept it more. However, if it is too hyper-realistic, people may find it creepy [28]. This is also supported by [29] stating that “hyperrealistic-animated appearance can make people experience more feelings of eeriness than those with a cartoonish-still appearance, which can diminish the positive effects of high anthropomorphism on warmth perceptions.” It is recommended to give it a cute baby-faced feature instead.

2.4.2 AI Musician Capability

AI Musician Capability as a concept refers to the sources considering an AI musician’s capabilities at any point of the study. This general term may refer to various different things such as an AI’s competency, autonomy, scalability, efficiency, cost, music quality and any other indicators of an AI’s capabilities with regards to music. AI musicians may be much more scalable, can create music faster and cost less. [29] However, AI autonomy is also a point to consider [25] as creativity is typically dependent on the programmer. Autonomy refers to the degree in which the AI depends on a human to provide input/feedback in order to compose/perform. Through the study of [25] it was concluded that it does not affect acceptance of AI.

Based on the sources, in general the quality of AI music is lower than that of human music. This is so as in the empirical studies [25, 29, 30, 33] where participants of the study are not informed that the composer of a music piece is an AI, the piece is still rated lower compared to that of a human composer, leading to the possibility of AIs producing lower quality music. However there are other several factors that may play a role in this as the researcher’s/studies’ music piece selection may play a crucial role, along with participant’s personal music tastes. Perhaps the AI music pieces chosen were simply not as good as the one chosen for human composers.

Aside from musical composition, AI musicians are able to do many things human musicians are not able to in terms of performance such as instant outfit changes, appear and disappear at any moment and perform stunts/choreography humans may not be physically capable of.[28]

2.4.3 Prejudice/Stereotype

Prejudice/Stereotype Towards AI Musicians refers to the concept of humans having a misconception or strong opinion towards what an AI should be and their characteristics/capabilities. This may cause humans to prefer human musicians and have a negative response towards AI musicians not because of their capabilities but simply due to being non-human.

According to [25], humans may question an AI’s creativity. People see them not as the creative type as AIs require a human programmer to program it for them. This may cause humans to not acknowledge AIs as musicians due to role theory. In the social status of individuals, AI’s role is to substitute human labor. This is due to the way they create songs. This is very well explained by [25]. “ If a machine does not make music autonomously with its own creativity, it means that the machine relies on human inputs in order to be creative. In this case, people will see this machine less as a musician but rather as a musical instrument, such as software that electronic dance music (EDM) musicians use.” In summary, if an AI is used more like a tool rather than being an autonomous composer who does not require human intervention, they will not be seen as musicians.

This prejudice is especially prevalent in traditional music. Such is the case in Irish traditional music [33]. This is so as traditional music tends to be very unconventional to the normal genres and may be unique to a region. In the study of [33], ITM(Irish Traditional Music) practitioners rank pieces they think belong to AI very harshly compared to those they think are human. ”This difference points to a potential conscious prejudice regarding what a subject believes the capacities of AI or computers are compared to humans when it comes to ITM composition.” [33]

2.4.4 Biased Towards Human

The concept of *Biased Towards Humans* is similar to the previous concept in terms of having humans prefer human musicians over AI. However, the reason for this preference is different. This concept refers to humans preferring other humans over AI neither because of discrimination nor prejudice towards AI but simply due to the fact that some may value human work better due to the cultural/emotional value it may bring.[30] This is prominent in the empirical study [30] where informing respondents of the nature of the composers resulted in a statistically significant and positive

effect on the evaluation of human-generated music compared to music generated by the AI.

2.4.5 Fear Towards AI

The concept of *Fear Towards AI* refers to humans preferring to not support AI not due to preference of humans but due to disliking/fear of AIs. This may come in various ways such as the fear of substitution or fear due to human mortality compared to AIs, which leans more towards human psychology.

Humans fear that AI may be their substitute and that they may lose their jobs/occupation due to their introduction & development thus their unwillingness to support AI.[30] This fear is valid for the context of music as the application of AI already has replaced several less creative jobs and if AI is to continue its development in a rapid pace, AI creativity may tend to improve.

Humans may also fear AI due to the fear of technological singularity and extinction. Humans are reminded of mortality as AI musicians are immortal as long as they are maintained. [28] This concept is related to Anthropomorphism, more specifically the uncanny valley theorem in which something that is too human-like but is not human is “perceived as a psychological discomfort similar to the experience of cognitive dissonance.” This reminds humans of self-preservation instincts and fear of extinction. [28, 29]

2.4.6 AI Acceptability and AI Music Evaluation

The last concept of Relation Between AI Acceptability and AI Music Evaluation refers to the possibility that a human’s perception, bias and preferences on AI may cause one to evaluate AI generated music or AI generated performances differently. All 5 sources mentioned this concept in one form or another. The sources suggest that there exists a positive correlation between AI acceptability and how humans may perceive and evaluate AI music. Meaning that the more an individual accepts the use/existence of AI, the more likely they are to rate AI music higher. This conclusion was made by several sources through statistical analysis of empirical studies [25, 29, 30, 33]. However, AI acceptance is not the only factor which may have an impact on an individual's perception of AI music and AI acceptance may also be affected by several other factors.

2.5 Analysis

Aside from acceptance, an individual’s evaluation of Artificial Intelligence music may be a result of several different factors. It may shift based on an individual’s warmth and competence perception (judgment of capability and relatability) of an AI [29, 33], an individual’s views of an AI’s role in society [25, 30], and the degree of which an AI is human-like [25, 28, 29]. It may also differ per individual depending on their psychological stance towards AI and their preference for a human touch[28, 30]. Another factor that evidently impacts AI music evaluation is the quality of the AI generated music itself. It is observed that AI generated music is generally rated lower in evaluations compared to human composed music without the disclosure of the music’s composer. This observation is further supported by the results of music expert evaluations. [30] However, there are exceptions to this, such as the case of Irish Traditional Music practitioners in which results show that generally, practitioners believe that a good ITM tune is not generated by AI. This is thought to cause the intentional poor evaluation of tunes thought to be generated by AI. [33] All these other factors can be categorized as an individual’s perception of AI musicians.

Some identified constructs are somewhat similar to one another and may perhaps be generalized. According to [47], prejudice and stereotypes are a source of bias within an individual, shifting one’s perception of certain things. With regards to this, it can be stated that prejudice and stereotypes may be a cause of bias. Thus, in order to generalize the 5 constructs of Anthropomorphism, Fear (towards AI), Capabilities (of AI musician), Stereotype/Prejudice (against AI), and bias towards humans, the last two concepts may be considered as Bias due to their intertwined relationship and definition. This suggests that the 4 constructs of *Anthropomorphism, Fear, Capabilities* and *Bias* may be used as a measure for one’s perception of AI musicians and performers. In a later section, this study will propose a scale in order to measure one’s perception of AI musicians and performers using these 4 constructs.

According to the observations made from the systematic literature review, some conclusions can be made. Firstly, an individual’s evaluation of AI music may be influenced by one’s acceptance level of AI. This is based on several sources through statistical analysis of empirical studies [25, 29, 30, 33] in which AI acceptance is present in one form or another. The sources suggest that there exists a positive correlation between AI acceptability and how humans may perceive and evaluate AI music. Meaning that the more an individual accepts the use/existence of AI, the more likely they are to rate AI music higher. This may be the case as individuals who are more open towards the use of AI may harbor a higher view of AI in the social hierarchy.

This leads towards the second conclusion that an individual’s evaluation of AI music may be influenced by one’s perception of AI musicians/performers. One’s perception of AI musicians may be influenced by several factors such as an AI’s degree of Anthropomorphism, AI’s capabilities, quality of music and several other factors. It is expected that the greater one’s perception of AI musicians/performers, the greater one will evaluate AI music. However, there may be exceptions to this especially in cases in which the perception of AI musicians is exceedingly high. In these cases, an individual’s expectation

towards the AI musician becomes too high that the music is evaluated very harshly due to the rise in standards & expectations. [29]

Lastly, it is concluded that AI generated music is of lower quality than that of human composed music as it seems to be the case throughout the sources. However, a point to consider is the researcher's selection of music to use in their respective studies. It is possible that the musical piece(s) chosen for the studies may not have been equal in all aspects resulting in a discrepancy in quality between the musical pieces being compared. This however is difficult to evaluate due to subjectivity, personal tastes and preferences. The age of the papers may also have an effect as the development of AI is continuing to improve. As the study uses sources dating back to 2021 in the literature review, there is a possibility that the quality of AI music has changed since then.

2.6 Hypothesis

A quantitative study is designed based on these statements in order to investigate the accuracy of these relationships and further support or disprove these statements. It is expected that during music evaluation, participants will rate the music piece composed by AI lower than that of the human composer without knowing the nature of the composer of either music piece. It is also expected that participants will think that the lower rated music piece is the one created by an AI when asked. Finally, it is also expected that the AI acceptance scale and perception of the AI musician scale will have a positive correlation to the music evaluation of the AI composed piece.

In accordance to the findings & observations of the literature study and the sources, the following hypotheses are formed:

H1. An individual's level of AI acceptance has a positive correlation with the individual's perception towards AI musicians/performers.

H2. An individual's level of AI acceptance has a positive correlation with the individual's evaluation of AI music.

H3. An individual's perception towards AI musicians/performers has a positive correlation with the individual's evaluation of AI music.

H4. Music affiliated to AI musicians/performers are of lower quality compared to music affiliated to humans. (AI music will be evaluated lower compared to human music)

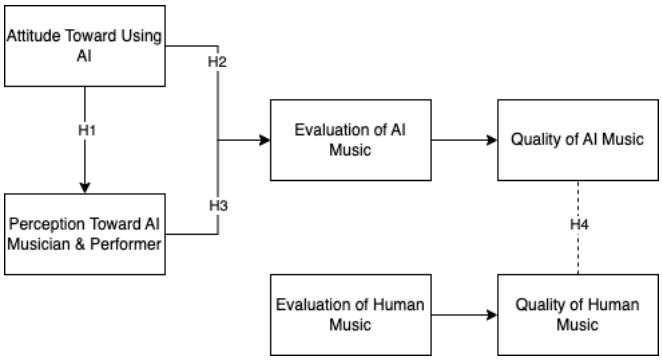


Figure 3. Hypothesis Conceptual Model

These hypotheses were formulated with the assumption that AI acceptability influences one's perception of AI musicians, and that AI acceptability and perception towards AI musicians influences one's evaluation of AI music, but not vice versa. This is the shared understanding derived from the sources. Depending on the results of these hypotheses, it would suggest or confirm the presence of bias in the perception of AI musicians/performers and possibly elaborate on the source(s) of bias. In order to conduct the study to confirm or deny these hypotheses, scales to measure certain concepts must be selected. Scales will be needed to measure an individual's AI acceptability and perception towards AI musicians/performers.

2.7 Measures

With relevance to H1 and H2, AI acceptability will be measured through the use of an established and validated *unified theory of acceptance and use of technology* (UTAUT) scale. The items on the scale may vary depending on which specific scale will be used but generally UTAUT scales include items to measure performance expectancy, effort expectancy, attitude towards using technology, social influence, facilitating conditions, self-efficacy, anxiety, and behavioral intention to use the system. [34] However, due to the length of the survey, the *technology acceptance model* (TAM) [42, 43] is used in its place. A TAM model only includes items to measure perceived ease of use and perceived usefulness but contains a considerably lower amount of items. Although TAM may not provide insight as significant as the UTAUT scale, TAM is a scale that will be able to provide a satisfactory and validated measure of how much an individual accepts AI. This scale will be adopted from another study [42].

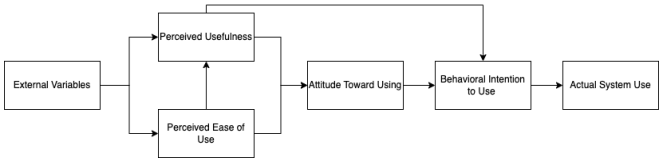


Figure 4. TAM Conceptual Model

As of the writing of this paper, validated and established scales to measure the general perception of AI musicians and performers are very limited. Thus to investigate H1 & H2, the study will develop a scale based on the concepts and results of the literature study. The scale will include items pertaining to Anthropomorphism of AI musicians, AI musician capabilities, fear towards AI and bias (role of AI, stereotype and prejudice). As a scale will be developed, it will need to be validated to ensure it is a reliable scale.

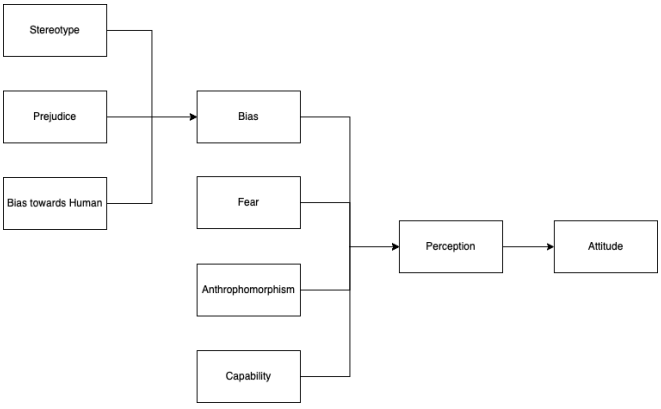


Figure 5. Conceptual Model

The scale for music evaluation which is relevant to H2, H3 & H4 will be adopted from [25]. This scale contains nine likert-scale items and has also been used in previous studies [36] as a scale for the assessment of musical quality, thus this scale is deemed reliable. The scale was initially developed from “Rubric for assessing general criteria in a composition assignment” [35]. The main components for music evaluation in this scale includes aesthetic appeal, creativity and craftsmanship.[36] In this scale, a greater score indicates a more positive evaluation of musical quality. This scale will be used to evaluate both AI and human music in order to gauge their quality.

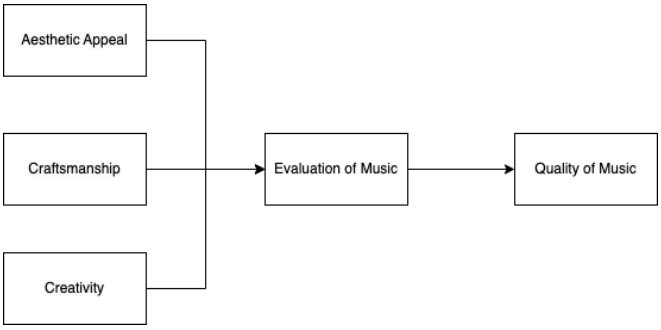


Figure 6. Evaluation of Music Conceptual Model

3 METHODOLOGY

A quantitative study will be performed in order to investigate the relationship between AI acceptability and one’s perception towards AI musicians/performers. The qualitative study will be conducted in the form of a Qualtrics Survey [21].

3.1 Population & Demographics

With the purpose of limiting the number of variables that may influence the study, the participants for this survey will be restricted to individuals from or formerly from the University of Twente. This target population was chosen due to familiarity with the target population and as it is believed that most individuals in this group may have basic knowledge of Artificial Intelligence. The participants' age, level of study, gender and other information will be recorded as demographic information in order to find possible abnormalities or similarities in certain results with regard to specific demographic items.

In order to perform data analysis, a recommended minimum of 5 participants per item in the scale is required when accessing a scale. [20] Thus, the minimum number of participants for the survey will be the product of 5 and the number of items present in the scale.

3.2 Scales & Considerations

The order of blocks in the survey are as follows:

1. Informed Consent
2. Demographic
3. Music Evaluation
4. AI Acceptability Scale
5. Perception Towards AI Musician/Performer Scale

The order of measurements in the survey was intentional and planned. The first block to be seen by participants will be the Informed Consent block. The block informs the respondent about the study and use of their data, thus giving them a choice to either participate or not participate in the study. Afterwards, the demographic items were placed with the intention of gathering participant’s data that may still be used in the case that the respondent does not complete the survey.

The next three blocks are then placed in this order to reduce the possibility of participant’s answers being influenced by previous scales/items. All the items in this study were measured using a 5 point likert-scale. The music evaluation block will present participants with two audio files of a similar genre of music. One is generated by AI and one composed by a human musician. A random half of the participants will be informed of the music composer’s nature. The participants must evaluate the quality of the two pieces. This block is placed above the scales due to its significance. To minimize

the possibility of bias in this block, the order of questions and whether or not the participant is informed of the composer's nature is randomized. The scale for this block was adopted from [25]. Higher scores indicate a greater evaluation of the musical piece. This block is necessary to validate or reject H2, H3 and H4 by measuring one's evaluation of music.

An AI acceptability scale is placed in the next block to measure a participant's AI acceptability without influencing the classification. This block uses a TAM model adopted from [42]. Higher scores indicate greater acceptance of AI. The final block is the Perception of AI Musician scale. This block was placed last due to its complexity and is expected to contain the most items. Thus, there may be a possibility that participants may not continue the survey if this block was placed before the others. This block contains 4 subscales based on the 5 concepts mentioned in section 3.4 with a high score indicating positive perception of AI musicians/performers. Higher scores indicate a higher perception of AI musicians/performers. An open question was also included at the end of the block for respondents to express their views of AI musicians and performers. Aside from the order of blocks, the consideration was also made to give definitions and examples of uncommon terms and difficult questions along with some images to further support the scales for better participant understanding. (*See Appendix B*)

3.3 Procedure

Before the survey is officially published, a pilot testing of the survey will be conducted on a small group of University of Twente students with the goal of ensuring that the survey was not too difficult and to estimate the time it will take to complete the survey. Several adjustments may need to be made based on the response and comments of the pilot. The two musical pieces chosen for the survey were of the classical genre. Both pieces were piano solos with similar pace and length. The classical genre was chosen due to the great availability of classical AI music. The survey was then published and distributed through various online means to University of Twente students. This may include various discord groups, WhatsApp groups, MicrosoftTeams chats and distributed through personal chat to students. Aside from online distribution of the survey, if the need arises, the survey will also be distributed physically in University of Twente grounds by distributing QR-codes of the survey to students in public areas while offering sweets, candies or snacks in return.

The survey received 122 respondents in which 64% were male and 36% were female. The average age of respondents was 27 with 58% of respondents originating from Europe, 40% from Asia, 1% from South America and 1% from Oceania. From a scale of 1 to 10 with 10 being the most, respondents on average scored 7.3 on frequency of listening to music and 2.9 on practicing music. This would suggest that on average, the respondents are familiar with music to a certain degree but may not be familiar with the creation/playing of music and musical instruments. After any respondents' personal information and missing responses were removed, 94 of the 122 responses were deemed usable for the study.

3.4 Validation

To investigate the validity of the measurement items, constructs and scales, factor analysis and Cronbach's α test was conducted. The evaluation of music (EoM) scale and all constructs of the TAM used as the AI acceptability scale had Cronbach's $\alpha > 0.8$ indicating great internal consistency of the scales. (*See appendix A*) This was to be expected as these scales have been validated in previous studies. The constructs of bias and anthropomorphism within the perception of AI musician and performer scale (PREAMP) failed to obtain the recommended Cronbach's α coefficient within the range of 0.7 to 0.9 [44], instead obtaining an $\alpha < 0.7$. The factor loadings of the two constructs revealed that two items may potentially be the cause of validity issues due to their incredibly low factor loading of < 0.1. Items A1 and B1 were therefore removed from the factor analysis and will no longer be included in further calculations within the study. The removal of items A1 and P1 resulted in all constructs and the full scale fulfilling the threshold $\alpha > 0.7$, verifying the internal consistency and reliability of the items to be used as factors.

Construct	No.	Factor Loadings	Initial Cronbach's Alpha	Cronbach's Alpha	Average Score
Bias	B1	-0.024	0.636	0.717	2.690
	B2	0.302			
	B3	0.747			
	B4	0.978			
	B5	0.492			
Anthropomorphism	A1	0.091	0.591	0.704	2.742
	A2	0.534			
	A3	0.923			
	A4	0.587			
Fear	F1	0.690	0.771	0.771	3.440
	F2	0.783			
	F3	0.711			
Capability	C1	0.790	0.772	0.772	3.333
	C2	0.823			
	C3	0.637			
	C4	0.446			
	C5	0.572			
	C6	0.403			
Full Scale			0.739	0.750	3.035

Table 1. Factor Loadings & Cronbach's Alpha

4 RESULTS

To test H4, which predicts that human music will be scored more highly compared to AI music, and observe whether informing participants of the nature of the music composer (AI or human) has an impact on the evaluation of music, the responses from the evaluation of music scale were tested and compared using three two-sample t-tests. (*See table 3*) The results of the t-test showed that the difference in mean

scores, whether the respondent is informed or not informed of which music piece was generated by AI, shows no statistical significance (at a 95% significance level) in music evaluation scores for both human and AI composed music. Meaning that a participant's knowledge of whether a music was AI generated or human made has no effect on either music piece's evaluation However, the difference in mean scores for AI music and human music is statistically significant at a 95% significance level. The result of the t-test supports H4 as the mean score of human music was greater than that of AI music in the instance of this t-test.

Relation	Pearsons Correlation Coeficient	p-value	Significance
AIA & EoM	0.397	9.57062E-05	**
AIA & Bias	0.101	0.3399	-
AIA & Anthropomorphism	0.129	0.2246	-
AIA & Fear	-0.060	0.5743	-
AIA & Capability	0.310	0.0028	**
AIA & AI Musician Scale	0.206	0.0497	*
EoM & AI Musician Scale	0.288	0.0057	**
EoM & Bias	0.171	0.1054	-
EoM & Anthropomorphism	0.220	0.0360	*
EoM & Fear	-0.116	0.2715	-
EoM & Capability	0.405	6.97898E-05	**
PU & EoM	0.356	0.0005	**
PU & AI Musician Scale	0.354	0.0006	**
PEoU & EoM	0.323	0.0018	**
PEoU & AI Musician Scale	0.123	0.9023	-

Table 2 .Pearson's Correlation
*Note: *Correlation is significant at the 95% level.*
***Correlation is significant at the 99% level.*

Pearson's correlation coefficient for pairs of variables (scales and concepts/constructs) and the statistical significance of the coefficients were calculated to test H1, H2 and H3. The variables tested were AI acceptance (AIA), AI music evaluation score (EoM) and the perception of AI musician/performers (PREAMP) scale and its constructs. Out of the 11 coefficients calculated, only 6 were statistically significant.

H1 predicts that AI acceptance and perception towards AI musicians/performers are positively correlated. Correlation coefficients show a weak positive correlation (0.206) between AIA and the perception of AI musician and performer (PREAMP) scale at a 95% significance level supporting the hypothesis.

H2 predicts that AI acceptance and music evaluation score are positively correlated. Correlation coefficients show a slight positive correlation (0.397) between AIA and the evaluation of AI music score at a 99% significance level supporting the hypothesis.

H3 predicts that perception towards AI musicians/performers and music evaluation score is positively correlated. Correlation coefficients show a weak positive correlation (0.288) between evaluation of AI music score and the PREAMP scale at a 99% significance level supporting the hypothesis.

The open question asking for respondent's opinion of AI musicians/performers received 67 valid responses. 27% (18) of responses provided positive opinions, 30% (20) provided negative opinions and 43% (29) provided mixed/neutral opinions. Several trends may be observed throughout the responses.

- 14 out of 67 responses state that AI musicians/performers' lack human emotion or feeling.
- 13 out of 67 responses regard AI musicians as a tool used by humans.
- 11 out of 67 responses prefer human musicians and states that humans are better than AI and may not be replaced.
- 8 out of 67 responses state that AI musicians lack creativity and do not produce original music.
- 7 out of 67 responses state that AI musicians are simply copying or mimicking human music.
- 6 out of 67 responses view AI musicians as a threat to human musicians due to the possibility of substitution.
- 5 out of 67 responses view AI musicians as capable and may have hope in its development and innovation.
- 4 out of 67 responses state that AI musicians possess great potential for personalized music.
- 4 out of 67 responses state that AI musicians lack story in their music that humans obtain through experience.

The trends may not be proven to be statistically significant but may still provide valuable insight nonetheless. The open question does not give any predefined answer to respondents and respondents are free to enter any of their thoughts and opinions pertaining to AI musicians and performers, thus trends or similarity in answers may be useful to serve as a minor qualitative part for the study. This question provides some context and reasoning behind why the results pertaining to the relationships of constructs are as they are.

6 DISCUSSION

The results of music evaluation score (EoM) suggests that AI music is lower in quality compared to human music based on the fact that the difference in the two evaluation scores is statistically significant. According to observations from the open opinion question, the lower evaluation of AI music may be caused by the lack of emotion or feeling in the musical piece. This is so as emotion is an integral part of music and "good music does not just blindly follow rules, it has feeling,

emotion" and the current music generating AIs are lacking in the emotional component of music perception and production.[45] However, the results may not be an accurate depiction of the comparison between all AI music and human music quality. It is true that in this case two similar musical pieces of similar length, genre, style, pace and other traits were compared to each other, but not all musical pieces are viewed equally and different musical pieces are of different quality regardless of the nature of its composer. There is also the matter of musical taste and preferences of different individuals that may play a role in the evaluation of music affecting its assessment. [30]

Hypothesis	t-value
There is a difference in AI average score when informed and not	0.861
There is a difference in means between AI & Human average scores	0.035
There is a difference in Human average score when informed and not	0.592

Table 3. EoM t-test

An interesting finding from the results of the t-tests performed on the music evaluation scores is the fact that there is no statistical significance between the scores of AI nor human music evaluation when participants are informed of the composer's nature (being an AI) and when participants are not informed. This would suggest that knowing if a musical piece was created by an AI or human musician does not affect an individual's evaluation of music quality. This result was unexpected as previous studies suggest the opposite. According to [30], individuals will be in favor of human-made alternatives rather than AI-made in regards to creative products based on an empirical study, using music, with some similar aspects to this one. However, this was not reflected in the current study. Results from the open opinion question suggest that this is possibly due to individuals being open to the idea of AI music as long as the music is "good". This difference may also possibly be due to the study's research population being members of a technical university in which many may be familiar with the use of AI making them more open to the idea.

Relation	Pearsons Correlation Coeficient	p-value	Significance
Bias & Anthropomorphism	0.298	0.0042	**
Bias & Fear	0.156	0.1400	-
Bias & Capability	0.430	2.12296E-05	**
Anthropomorphism & Fear	-0.080	0.5507	-
Anthropomorphism & Capability	0.226	0.0312	*
Fear & Capability	-0.080	0.4524	-

Table 4. Construct Correlation

The correlation between constructs in PREAMP supports findings of previous studies [25, 29]. Based on the results, bias and anthropomorphism, bias and capability, and anthropomorphism and capability had statistically significant positive correlations between constructs. A higher level of anthropomorphism leads individuals into viewing AI musicians as more competent [29] or in this case, more capable. This increase in perception of capability in turn results in individuals to be more accepting of an AI's musician role [25], possibly reducing the prejudice and stereotyping which builds into the bias construct, explaining all three correlations.

Aside from the correlation of PREAMP among constructs, six correlations were statistically significant (See *table 2*). Among them is the positive correlation between AIA & EoM which would suggest that one's level of AI acceptance would have an impact on one's evaluation of music. The results also suggest a positive correlation between AIA and capability, this is likely caused by the presence of perceived usefulness as part of the TAM used to measure AIA. Thus, the more capable/competent an AI is, the more useful it may be. AIA and the perception of AI musicians and performers have a positive correlation. As previously mentioned, perceived usefulness is a part of AIA. Perceived usefulness plays a mediating role on one's perception of AI [46], thus it would be expected that the more accepting one is of AI, the more likely they are to appreciate AI application in different fields including that of music.

The remaining three correlations of EoM and PREAMP, EoM and Anthropomorphism, and EoM and competence may have related explanations. Firstly, although the study did not present participants with an image or description of the AI musician, there is still a correlation between anthropomorphism and EoM. Perhaps respondents had a particular picture in mind of the composer when listening to the music but this result contradicts that of previous study [25] in which it was concluded that anthropomorphism has no effect on the assessment of AI-composed music. Capability and EoM having a positive correlation is to be expected and straightforward. The more capable and competent the AI musician is, the higher the quality of music generated by the AI musicians and a higher quality of music would result in a greater score in its evaluation. As for EoM and PREAMP, it would seem that the greater one's perception of an AI musician is, the more individuals may respect the AI musician resulting in a higher scoring. An example of this is how the work of a respected painter may be valued more than the work of a random person due to the known identity of the painter. Based on these correlations, it may also be considered that the two constructs of anthropomorphism and capability are the two most impactful constructs of PREAMP with regards to the evaluation of music.

6.1 Future Work & Limitations

This study creates a scale to measure an individual's perception towards AI musicians and performers (PREAMP) based on the information gathered in the systematic literature review from 5 sources. Although validated to a certain degree, further validation of its constructs and items may be required in order to improve and generalize the scale for other use within the AI creativity field such as art. However, in

order to generalize into other fields, another literature review is needed to adapt the scale to ensure that the constructs in the PREAMP scale may be used in that field of AI. Ideally items with factor loadings of < 0.5 (B1, B2, B5, A1, C4 & C6) are to be replaced by other items that may be better suited for their respective constructs. A factor analysis may also be conducted on the scale to measure the amount of factors that the scale is measuring to ensure that it is measuring the constructs as intended.

The sample population used in the quantitative study was students of the University of Twente which may not be an accurate depiction of the population. This is so due to the demographics age only covering a limited range, a majority of participants only coming from 2 continents, the familiarity and knowledge of AI which may affect participants' views and other factors which do not reflect the general population perfectly. A study could be conducted with participants categorized into different age groups, different backgrounds and other demographics in order to investigate their impact and provide a more accurate sample of the population. The selection of musical pieces for the evaluation may also be expanded. The current study only provided participants with musical pieces of the classical genre which may not be suitable for all participants, thus a study with multiple genres and musical piece selections may be conducted. The study investigates the relationships between perception, music evaluation and AI acceptance. Through a quantitative study, the relationships were supported and studied. However, the reason behind the results are not confirmed. Thus, a quantitative study may need to be conducted to explore the reasoning behind these relationships and why the relationships are as they are.

The study focuses on an individual’s perception towards AI musicians and performers and its influence on an individual’s perception of music. However, there are other areas in which this study does not take into account, namely; ethics, reliability, ownership, copyright and other areas. There exist cases in which the voices and songs of human artists and musicians were used in the training of AI musicians causing the AI to generate music with voices similar to that of a human musician. [48] These situations are quite delicate and are in the “gray area” as copyright laws for AI are still not fully developed due to its recent emergence. There is also the topic of responsibility and ownership regarding AI musicians in which it is unclear whether the programmer or the AI itself is to be held accountable in troublesome situations. AI musicians may also have the ability to replicate the music style of human musicians, possibly generating “fake” music of that human musician. Nevertheless, there are numerous areas in which this study does not fully take into account. Thus, future studies may take this into account and expand the PREMP scale to include additional constructs such as ethics.

7 CONCLUSION

The purpose of this study was to investigate how the perception of humans/music listeners towards artificial intelligence may influence the evaluation and perception of AI performers and AI musicians. A quantitative study was designed based on findings of the literature review in order to investigate the accuracy of relationships and support or disprove several statements. It is hypothesized that during music evaluation, participants will rate the music piece composed by AI lower than that of the human composer without knowing the nature of the composer of either music piece. It is also expected that one's level of AI acceptance may influence their perception of AI musicians/performers and how they evaluate AI music. Finally, it is also expected that one’s level of AI acceptance and perception of an AI musician/performer will have a positive correlation to the music evaluation of the AI composed piece.

According to the observations made from the systematic literature review and hypothesis testing via quantitative study, some conclusions may be made and all hypotheses were statistically supported. Firstly, it is concluded that currently the quality of AI music is lower than that of human music. This conclusion was drawn from t-tests of the evaluation of music. It is also concluded that AI acceptance indeed positively impacts the perception of AI musicians and performers due to the mediating role of perceived usefulness. And also that both perception and AI acceptance has a positive correlation to the evaluation of AI-composed music.

If the conclusion pertaining to quality is applicable not only for the context of AI musicians but instead Artificial Intelligence as a whole, it may help in addressing the skepticism towards the use of AI. Theoretically, if the quality of service, product or any output provided by AI is currently lacking compared to that provided by a human, it may rationalize the implications regarding trustworthiness and explainability as why choose AI if humans do it better. Perhaps as AI continues to advance, there will be a time where AI will be able to match or even surpass that in which a human may provide while also being more reliable. During that time, perhaps humans may learn to put more trust in AI and acknowledge them to the point in which the general consensus of AI has shifted to their favor. However, along with this many more ethical & human rights concerns may arise.

REFERENCES

[1] Cattell, R. B., & Anderson, J. C. (1953). The measurement of personality and behavior disorders by the I. P. A. T. Music Preference Test. Journal of Applied Psychology, 37(6), 446–454. https://doi.org/10.1037/h0056224

[2] Hernandez-Olivan, C., & Beltrán, J. R. (2022). Music Composition with Deep Learning: A Review. In Signals and communication technology (pp. 25–50). https://doi.org/10.1007/978-3-031-18444-4_2

[3] Kenmochi, H., & Ohshita, H. (2007). VOCALOID - Commercial singing synthesizer based on sample concatenation. Conference of the International Speech Communication Association, 4009–4010. http://www.let.uu.nl/~Gerrit.Bloothooft/personal/SSC/Yamaha/VOCALOID_Interspeech.pdf

[4] Moura, F. T., & Maw, C. (2021). Artificial intelligence became Beethoven: how do listeners and music professionals perceive artificially composed music? Journal of Consumer Marketing, 38(2), 137–146. https://doi.org/10.1108/jcm-02-2020-3671

[5] Petitbon, A. M., & Hitchcock, D. B. (2022). What kind of music do you like? A statistical analysis of music genre popularity over time. Journal of Data Science, 168–187. https://doi.org/10.6339/22-jds1040

[6] Shepherd, D., & Sigg, N. (2015). Music preference, social identity, and Self-Esteem. Music Perception, 32(5), 507–514. https://doi.org/10.1525/mp.2015.32.5.507

[7] Snowdon, C. T., Zimmermann, E., & Altenmüller, E. (2015). Music evolution and neuroscience. In Progress in Brain Research (pp. 17–34). https://doi.org/10.1016/bs.pbr.2014.11.019

[8] Upadhyay, D., Shukla, R., & Chakraborty, A. (2016). Factor structure of music preference scale and its relation to personality. Journal of Indian Academy of Applied Psychology, 43(1), 104-113.

[9] Zulić, H. (2019). How AI can change/improve/influence music composition, performance and education: three case studies. INSAM Journal of Contemporary Music, Art and Technology, 1(2), 100-114.

[10] Fitch, W. T. (2005). The evolution of music in comparative perspective. Annals of the New York Academy of Sciences, 1060(1), 29-49.

[11] Ashrafian, H. (2023). Artificial Intelligence as Art–What the Philosophy of Art can offer the understanding of AI and Consciousness.

[12] De Mantaras, R. L., & Arcos, J. L. (2002). AI and music: From composition to expressive performance. AI magazine, 23(3), 43-43.

[13] Shank, D. B., Stefanik, C., Stuhlsatz, C., Kacirek, K., & Belfi, A. M. (2023). AI composer bias: Listeners like music less when they think it was composed by an AI. Journal of Experimental Psychology: Applied, 29(3), 676.

[14] Lima, A. S., & Blixt, A. (2021). Investigating the possibility of bias against AI-computercomposed music. University of Skövde.

[15] Zlatkov, D., Ens, J., & Pasquier, P. (2023, April). Searching for Human Bias Against AI-Composed Music. In International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar) (pp. 308-323). Cham: Springer Nature Switzerland.

[16] Zenieris, R, (2023). Perception and bias towards AI music.

[17] Li, R., Yang, S., Ross, D. A., & Kanazawa, A. (2021). Ai choreographer: Music conditioned 3d dance generation with aist++. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 13401-13412).

[18] Wang, B., & Yang, Y. H. (2019, July). PerformanceNet: Score-to-audio music generation with multi-band convolutional residual network. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 1174-1181).

[19] Yamaha Artificial Intelligence (AI) Transforms a Dancer into a Pianist - News Releases - Yamaha Corporation. (n.d.). https://www.yamaha.com/en/news_release/2018/18013101/

[20] Hair, J. F., Tatham, R. L., Anderson, R. E., and Black, W. Multivariate Data Analysis, 6 ed. Prentice Hall, 2006

[21] Qualtrics XM - Experience Management Software. (2024, January 19). Qualtrics. https://www.qualtrics.com/

[22] Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. (2013). Using grounded theory as a method for rigorously reviewing literature. European Journal of Information Systems, 22(1), 45–55. https://doi.org/10.1057/ejis.2011.51

[23] Hanson, D., Storm, F., Huang, W., Krisciunas, V., Darrow, T., Brown, A. K., Lei, M., Aylett, M. P., Pickrell, A., & Robot, S. T. (2020). SophiaPop: Experiments in Human-AI Collaboration on Popular Music. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2011.10363

[24] Dong, J., Choi, K., Yu, S., Lee, Y., Kim, J., Vajir, D., Haines, C., Newbill, P., Wyatt, A., Upthegrove, T., & Jeon, M. (2023). A Child-Robot Musical Theater Afterschool Program for Promoting STEAM Education: A case Study and guidelines. International Journal of Human-Computer Interaction, 1–17. https://doi.org/10.1080/10447318.2023.2189814

[25] Hong, J., Fischer, K., Ha, Y., & Zeng, Y. (2022). Human, I wrote a song for you: An experiment testing the influence of machines' attributes on the AI-composed music evaluation. Computers in Human Behavior, 131, 107239. https://doi.org/10.1016/j.chb.2022.107239

[26] Lian, J. (2023). An artificial intelligence-based classifier for musical emotion expression in media education. PeerJ, 9, e1472. https://doi.org/10.7717/peerj-cs.1472

[27] Nautiyal, R., Jha, R. S., Kathuria, S., Chanti, Y., Rathor, N., & Gupta, M. (2023). Intersection of Artificial Intelligence (AI) in Entertainment Sector. 2023 4th International Conference on Smart Electronics and Communication .https://doi.org/10.1109/icosec58147.2023.10275976

[28] Sajadieh, S. (2023). Cute or creepy, that is the question of liveness: can artificial actors perform live? Artnodes, 0(32). https://doi.org/10.7238/artnodes.v0i32.412093

[29] Sun, D., Wang, H., & Xiong, J. (2023). Would you like to listen to my music, my friend? An experiment on AI musicians. International Journal of Human-Computer Interaction, 1–11. https://doi.org/10.1080/10447318.2023.2181872

[30] Tubadji, A., Huang, H., & Webber, D. J. (2021). Cultural proximity bias in AI-acceptability: The importance of being human. Technological Forecasting and Social Change, 173, 121100. https://doi.org/10.1016/j.techfore.2021.121100

[31] Yoon, S., & Kim, M. (2022). A Study on Deriving Improvements through User Recognition Analysis of Artificial Intelligence Speakers. Applied Sciences, 12(19), 9651. https://doi.org/10.3390/app12199651

[32] Liu, S., & Wang, P. (2023). Multi-dimensional fusion: transformer and GANs-based multimodal audiovisual perception robot for musical performance art. Frontiers in Neurorobotics, 17. https://doi.org/10.3389/fnbot.2023.1281944

[33] Déguernel, K., Sturm, B. L., & Maruri-Aguilar, H. (2022). Investigating the relationship between liking and belief in AI authorship in the context of Irish traditional music. HAL (Le Centre Pour La Communication Scientifique Directe). https://hal.archives-ouvertes.fr/hal-03849034

[34] Spil, A. A., & Schuring, R. (2006). The UTAUT questionnaire items. In IGI Global eBooks (pp. 93–98). https://doi.org/10.4018/978-1-59140-423-1.ch005

[35] Hickey, M. (1999). Assessment rubrics for music composition. Music Educators Journal, 85(4), 26–52. https://doi.org/10.2307/3399530

[36] Hong, J., Peng, Q., & Williams, D. (2020). Are you ready for artificial Mozart and Skrillex? An experiment testing expectancy violation theory and AI music. New Media & Society, 23(7), 1920–1935. https://doi.org/10.1177/1461444820925798

[37] Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial intelligence and human trust in healthcare: Focus on clinicians. Journal of Medical Internet Research, 22(6), e15154. https://doi.org/10.2196/15154

[38] Korteling, J., Van De Boer-Visschedijk, G. C., Blankendaal, R., Boonekamp, R., & Eikelboom, A. (2021). Human- versus Artificial Intelligence. Frontiers in Artificial Intelligence, 4. https://doi.org/10.3389/frai.2021.622364

[39] Ragot, M., Martin, N., & Cojean, S. (2020). AI-generated vs. Human Artworks. A Perception Bias Towards Artificial Intelligence? Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3334480.3382892

[40] Raso, F., Hilligoss, H., Krishnamurthy, V., Bavitz, C., & Levin, K. (2018). Artificial Intelligence & Human Rights: Opportunities & Risks. Social Science Research Network. https://doi.org/10.2139/ssrn.3259344

[41] Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in Human Resources Management: challenges and a path forward. California Management Review, 61(4), 15–42. https://doi.org/10.1177/0008125619867910

[42] Sohn, K., & Kwon, O. (2020). Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products. Telematics and Informatics, 47, 101324. https://doi.org/10.1016/j.tele.2019.101324

[43] Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. Management Science, 46(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

[44] Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. International Journal of Medical Education, 2, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

[45] Novelli, N., & Proksch, S. (2022). Am I (Deep) blue? Music-Making AI and emotional awareness. Frontiers in Neurorobotics, 16. https://doi.org/10.3389/fnbot.2022.897110

[46] Kim, J., Merrill, K., & Collins, C. (2021). AI as a friend or assistant: The mediating role of perceived usefulness in social AI vs. functional AI. Telematics and Informatics, 64, 101694. https://doi.org/10.1016/j.tele.2021.101694

[47] Jussim, L., Nelson, T. E., Manis, M., & Soffin, S. (1995). Prejudice, stereotypes, and labeling effects: Sources of bias in person perception. Journal of Personality and Social Psychology, 68(2), 228–246. https://doi.org/10.1037/0022-3514.68.2.228

[48] Times, N. (2024, May 10). Dutch MP's wants AI apps to compensate Dutch artists, stop using their songs, voices. NL Times. https://nltimes.nl/2024/05/10/dutch-mps-wants-ai-apps-compensate-dutch-artists-stop-using-songs-voices#:~:text=Research%20by%20AD%20found%20that,voices%20but%20without%20their%20permission.

APPENDIX

A APPENDIX

A.1 Factor loadings of AI musician/performer perception scale

Construct	No.	Factor Loadings	Initial Cronbach's Alpha	Cronbach's Alpha	Average Score
Bias	B1	-0.024	0.636	0.717	2.690
	B2	0.302			
	B3	0.747			
	B4	0.978			
	B5	0.492			
Anthropomorphism	A1	0.091	0.591	0.704	2.742
	A2	0.534			
	A3	0.923			
	A4	0.587			
Fear	F1	0.690	0.771	0.771	3.440
	F2	0.783			
	F3	0.711			
Capability	C1	0.790	0.772	0.772	3.333
	C2	0.823			
	C3	0.637			
	C4	0.446			
	C5	0.572			
	C6	0.403			
Full Scale			0.739	0.750	3.035

A.2 Average score of TAM model

Construct	Average Score	Cronbach's Alpha
Percieved Ease of Use	3.642	0.778
Percieved Usefulness	4.045	0.891

B APPENDIX

B.1 TAM Model used to measure AI acceptance. [42, 43]

TAM Construct	No.	Measurement Items	Scoring	Source
Perceived Ease of Use	PEoU1	Using the AI product would be easy	Normal	[42]
	PEoU2	Interaction with the AI product would be clear and understandable daily work performance	Normal	[42]
	PEoU3	I would find the AI product difficult to use	Reverse	[42]
	PEoU4	I would find it easy to get the AI product to do what I want it to do	Normal	[42]
Perceived Usefulness	PU1	Using the AI product would improve my daily work performance	Normal	[42]
	PU2	Using the AI product would help my daily work	Normal	[42]
	PU3	Using the AI product would enhance effectiveness in my daily work	Normal	[42]
	PU4	I would find the AI product useful in my daily work	Normal	[42]

B.2 Scale for Evaluation of Music [25]

Construct	No.	Measurement Items	Source
Evaluation of Music	EoM1	Many listeners would enjoy this music piece.	[25]
	EoM2	This music piece keeps listeners interested.	[25]
	EoM3	This music piece presented a strong aesthetic appeal.	[25]
	EoM4	This music piece was creative.	[25]
	EoM5	The music piece included very original musical idea (range, dynamics, timbre, tempo texture, rh	[25]
	EoM6	The music piece included unusual imaginative musical idea.	[25]
	EoM7	This music piece had a clear beginning, middle, and end.	[25]
	EoM8	This music piece appeared well-organized, not random.	[25]
	EoM9	This music piece had a good completeness overall.	[25]

B.2 Perception of AI Musician/Performer Scale

Construct	No.	Measurement Items	Scoring	Source
Bias	P1	AI music generators are just a tool used by human musicians.	Reverse	
	P2	AI music generators should not be recognized as musicians.	Reverse	
	P3	AI can never be as good as human performers.	Reverse	
	P4	AI can never be as good as human musicians.	Reverse	
	P5	AI are just copying human musicians.	Reverse	
Anthropomorphism	A1	AI performers give a sense of eeriness.	Reverse	
	A2	I would enjoy watching an AI performer.	Normal	
	A3	The human-like appearance of an AI performer makes me feel at ease.	Normal	
	A4	The human-like appearance of an AI performer makes me think they are capable.	Normal	
Fear	F1	I fear that AI would replace human musicians.	Reverse	
	F2	The fact that an AI musician has an endless lifespan worries me.	Reverse	
	F3	I fear the development of AI creativity.	Reverse	
Capability	C1	I think AI musicians are capable.	Normal	[29]
	C2	I think AI musicians are competent.	Normal	[29]
	C3	I think AI musicians are intelligent.	Normal	[29]
	C4	I think AI musicians are efficient.	Normal	[29]
	C5	I think AI musicians are skillful.	Normal	[29]
	C6	I think AI performers are better than humans due to their inhuman capabilities.	Normal	

C APPENDIX

C.1 Hypothesis Testing Pearson's Correlation

Relation	Pearsons Correlation Coeficient	p-value	Significance
AIA & EoM	0.397	9.57062E-05	**
AIA & Bias	0.101	0.3399	-
AIA & Anthropomorphism	0.129	0.2246	-
AIA & Fear	-0.060	0.5743	-
AIA & Capability	0.310	0.0028	**
AIA & AI Musician Scale	0.206	0.0497	*
EoM & AI Musician Scale	0.288	0.0057	**
EoM & Bias	0.171	0.1054	-
EoM & Anthropomorphism	0.220	0.0360	*
EoM & Fear	-0.116	0.2715	-
EoM & Capability	0.405	6.97898E-05	**
PU & EoM	0.356	0.0005	**
PU & AI Musician Scale	0.354	0.0006	**
PEoU & EoM	0.323	0.0018	**
PEoU & AI Musician Scale	0.123	0.9023	-

C.2 Music evaluation scale T-test (two-tail)

Hypothesis	t-value
There is a difference in AI average score when informed and not	0.861
There is a difference in means between AI & Human average scores	0.035
There is a difference in Human average score when informed and not	0.592

C.3 Music evaluation scale results

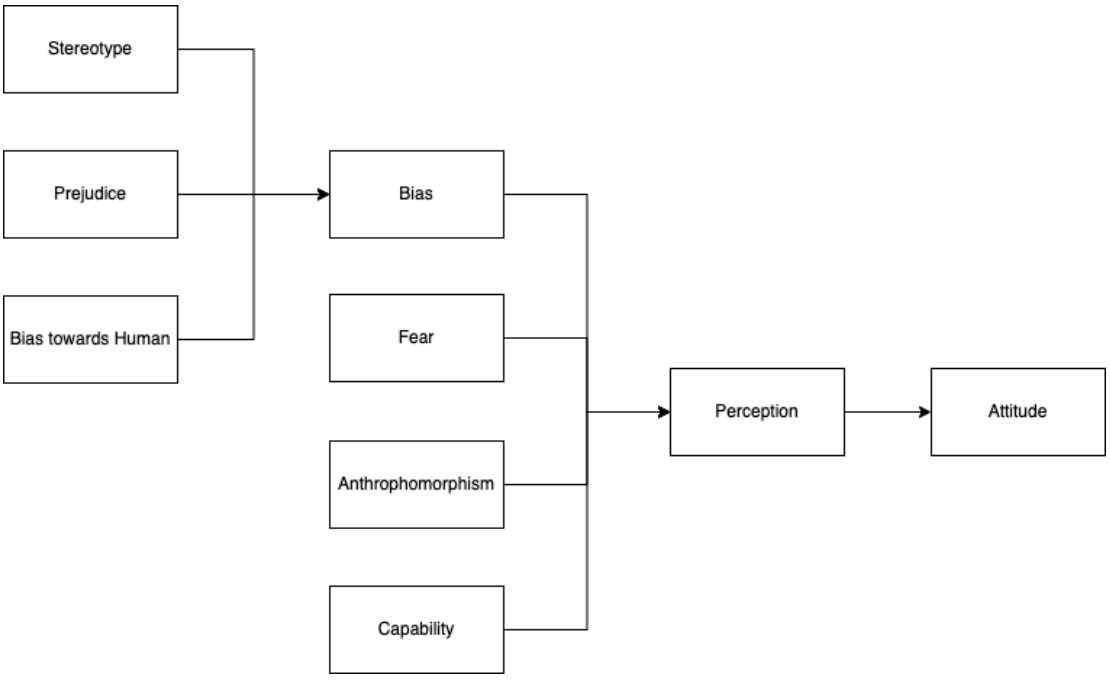
Item	Average Total Score	Average Score	Number of Respondents	Cronbach's Alpha
(1) AI Music	31.804	3.534	46	0.826
(2) Human Music	33.649	3.739	94	0.889
(3) AI Music - Informed	31.583	3.509	48	0.821
(1) & (3) AI Music Total	31.691	3.521	94	

C.4 Concepts Pearon's Correlation

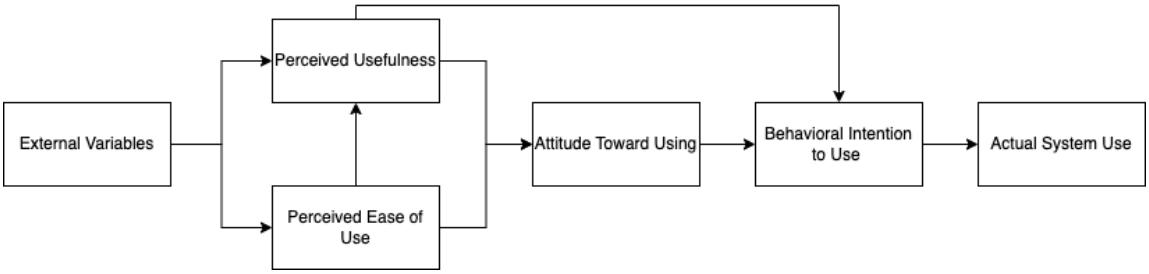
Relation	Pearsons Correlation Coeficient	p-value	Significance
Bias & Anthropomorphism	0.298	0.0042	**
Bias & Fear	0.156	0.1400	-
Bias & Capability	0.430	2.12296E-05	**
Anthropomorphism & Fear	-0.080	0.5507	-
Anthropomorphism & Capability	0.226	0.0312	*
Fear & Capability	-0.080	0.4524	-

D APPENDIX

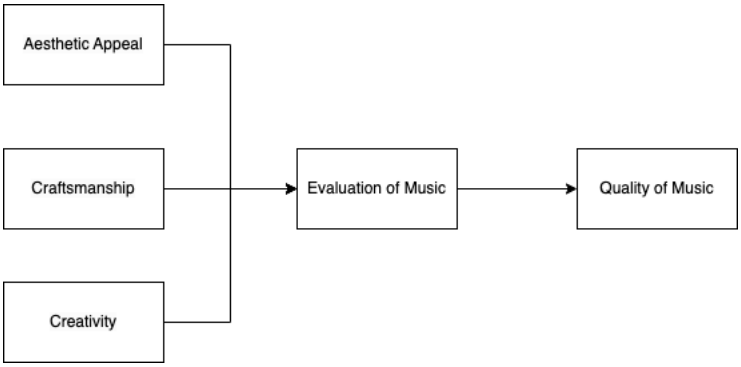
D.1 PREAMP Factor Relationship



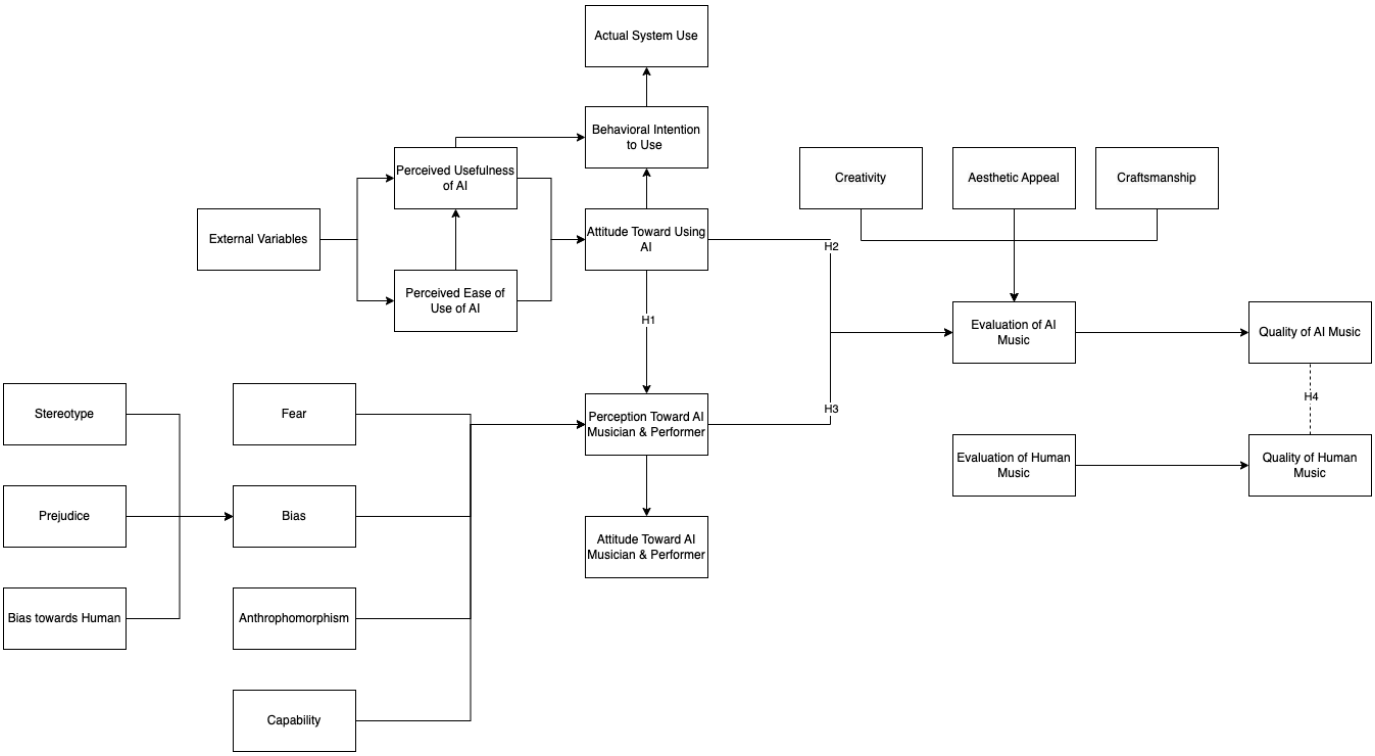
D.2 TAM Factor Relationship



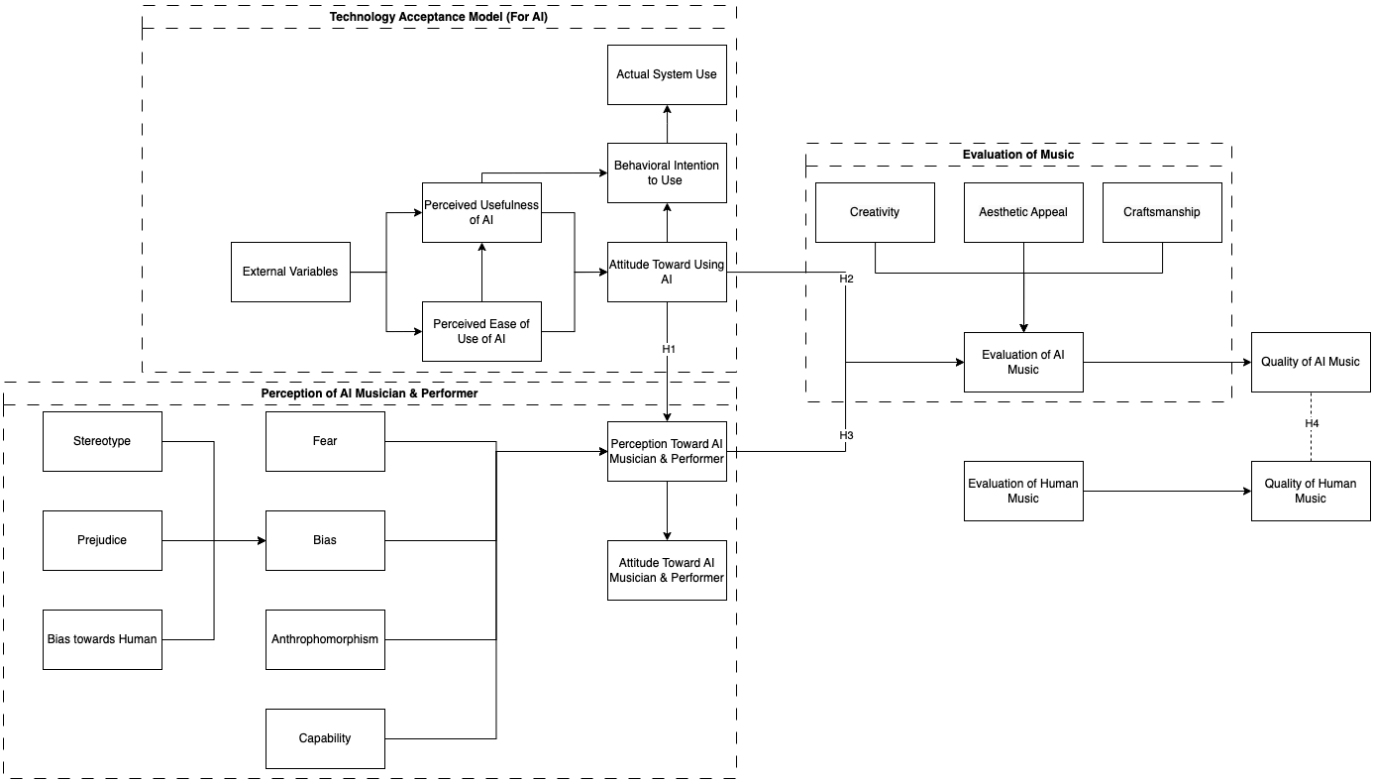
D.3 EoM Factor Relationship



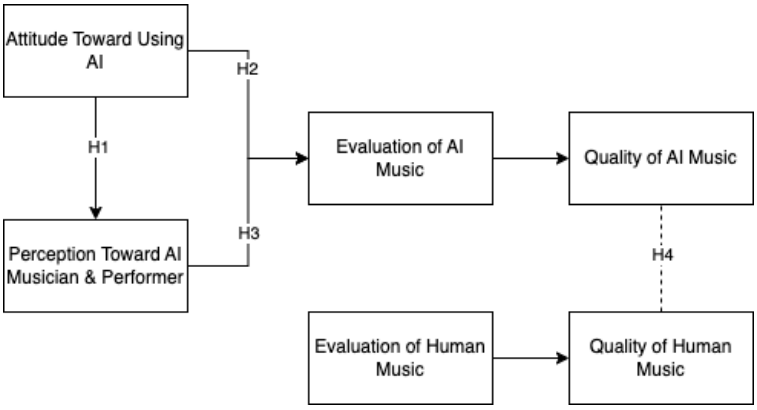
D.4 Complete Factor Relationship



D.5 Complete Factor Relationship 2.0



D.6 Simple Factor Relationship



D.7 Input Output Diagram