



Universiteit
Leiden
The Netherlands

Opleiding DSAI

Once upon a time
in a land far far away,
there was gender equality

Exploring gender stereotypes
in ChiSCor's fantasy stories told by children.

Kiara Titia Weterings
s2344114

Supervisors:

Dr. M.J. van Duijn & B.M.A. van Dijk

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

23/01/2024

Abstract

This thesis studies the gender stereotypes present in narratives of Dutch children. The literature review explores existing studies on gender stereotypes in early childhood. Two key studies are discussed, revealing insights into gendered intelligence stereotypes and gendered emotion stereotypes. This thesis uses a sample of 481 fantasy stories from the ChiSCor Corpus, and manual content analysis was used for coding character descriptions into categories. Statistical analysis of the 1536 characters mentioned in the sample reveals a gender imbalance, with more male characters on average. Significant associations are found between gender and emotional, intelligence, appearance, and personality descriptions. The findings confirm that the stereotype that females are more emotional is present in children's narratives. Contrary to the "brilliance is male" stereotype, the results indicate that female characters are more likely to have a positive intelligence description than male characters. In conclusion, the study highlights significant differences in how children describe genders in narratives, shedding light on the presence of gender stereotypes internalized by Dutch children. We advocate for further research so early interventions to challenge harmful stereotypes and promote diversity in children's narratives can be done.

Contents

1	Introduction	1
1.1	Problem statement	1
1.2	Research question	1
1.3	Thesis outline	2
2	Background literature	3
2.1	Presence of Stereotypes in Children	3
2.2	Gendered Intelligence Stereotype	3
2.3	Gendered Emotion Stereotype	3
2.4	ChiSCor	4
2.5	Definitions	4
3	Methods	5
3.1	Sampling methods	5
3.2	Content Analysis	5
3.3	Annotation Scheme	5
3.4	Statistical analysis	6
3.4.1	Average number of female vs male characters per story	7
3.4.2	Emotional Descriptions by Gender	7
3.4.3	Intelligence Descriptions by Gender	7
3.4.4	Appearance Descriptions by Gender	7
3.4.5	Ability Descriptions by Gender	8
3.4.6	Personality Descriptions by Gender	8
3.4.7	Occupations	8
4	Results	9
4.1	Average number of female vs male characters per story	9
4.2	Emotional Descriptions by Gender	9
4.3	Intelligence Descriptions by Gender	10
4.4	Appearance Descriptions by Gender	11
4.5	Ability Descriptions by Gender	12
4.6	Personality Description by Gender	14
4.7	Occupations	14
5	Discussion	18
6	Conclusion	20
	References	21
	Appendix	22
.1	Annotation Scheme	22

1 Introduction

Bian et al. (2017) discovered that when girls commence primary school at the age of five, they are equally inclined to believe that women could be 'really, really smart', as boys are to believe men can be 'really, really smart'. But when they turn six, something changes. Girls start doubting their gender. They even start adjusting their behavior. If a game is presented to girls as for 'children who are really, really smart', five-year-old girls are equally likely to want to play the game as boys. But after being in primary school for a year, six-year-old girls are suddenly uninterested while six-year-old boys are still interested [1]. This suggests that school teaches these girls a stereotype that brilliance is gendered and that this stereotype has an immediate effect on children's interests and actions.

1.1 Problem statement

Despite the findings by Bian et al. (2017) indicating that gender biases emerge among primary school-aged children, there is a lack of recent, natural context research that studies the nature and extent of gender stereotypes internalized by Dutch children. The current lack of an overarching assessment hinders our ability to understand and counteract the biases that shape the perceptions and behaviors of Dutch primary school children.

This is a problem because gender inequality in the Netherlands has not been solved. An example of this can be seen in She Figures (2021), which is a report that provides data on gender equality in research and innovation [2]. One of the things it shows is that while the proportion of females among overall doctoral graduates in the Netherlands in 2018 was 48.11%, which is a good increase since 2010 when it was 42.05, the proportion of females among doctoral graduates in 2018 differs across fields of study. In some fields of study, there are more women among the doctoral graduates, 68.3% women among doctoral graduates in Education, 62.0% in social sciences, and 59.3% in health and welfare. In other fields of study women are a minority such as 37.50% in Natural sciences, Mathematics and Statistics, 25.36% in Engineering, Manufacturing and Construction, and only 13.73% in ICT. This indicates that women are underrepresented among doctoral graduates in these STEM disciplines.

1.2 Research question

This study will attempt to fill the knowledge gap by using the ChiSCor database to test how stereotypes are portrayed in male and female characters by children in Dutch primary schools. The ChiSCor corpus contains fantasy stories spontaneously made up and told by children 4-12 years old, collected during 2020-2023. The stories were narrated within a natural context. This minimizes the risk of biases that occur because of controlled environments and ensures a high ecological validity [3].

Therefore, the question that this thesis will investigate the following research question:

“What are the differences in the way male and female characters are described by children (in primary school) in the stories of the ChiSCor database, with a focus on the categories emotions, intelligence, physical appearance, ability and personality?”

1.3 Thesis outline

To answer this research question, the following steps can be read in this thesis. First, a short overview of current literature about stereotypes present in Dutch children is given in chapter two. And a list of what kind of descriptions fall under each category. In the third chapter, the methodology is presented. Which consists of a description of the making of the annotation scheme, manual content analysis, and the statistical analysis that is used to make inferences using the ChiSCor data. Chapter four includes the results of this study, which are discussed in chapter five. Finally, chapter six includes the conclusion of this study.

2 Background literature

This chapter gives an overview of the literature related to this research. First an overview of studies concerning the presence of gender stereotypes in early childhood. Then two studies are mentioned that found gender stereotypes present in children. Following this, a section about the ChiSCor corpus is provided, including an overview of a case study about the gender bias in the ChiSCor stories. Finally, definitions as applied in this thesis, of the categories are given.

2.1 Presence of Stereotypes in Children

King et al. (2021) provides a systematic overview of the presence of gender stereotypes in early childhood, particularly the ages 3-5 [4]. They found evidence that children in this age group are aware and able to apply gender stereotypes. However, these stereotypes were easily influenced by the cultural and social environment. This emphasizes the importance of researching the current status quo of the stereotypes that are present in children, so effective strategies to counter harmful stereotypes can be designed.

2.2 Gendered Intelligence Stereotype

The previously mentioned study, Bian et al. (2017) was performed in the USA [1]. It consisted of two parts. The first part consists of three tasks, all designed to assess how often children (ages 5-7) rate their own gender as brilliant, which they refer to as own-gender brilliance. They corrected for other cues of intelligence such as professional dress and attractiveness. They conclude that a significant difference between own-gender brilliance scores for boys and girls emerged at age 6. However, they did not find a significant correlation between girls' perceptions of school achievement and their perceptions of brilliance. From their findings they conclude that the "brilliance = males" stereotype may be recognized, and applied by, children from the age of 6.

In part two of the study, they examine to what extent children's gendered beliefs about brilliance shape their interests. This was done by measuring the interest of boys and girls (aged 6-7) for two games, one for "children who are really, really smart" and the other for "children who try really, really hard". From their results, they concluded that the "brilliance = male" stereotype starts to affect children's interests as soon as children learn it.

2.3 Gendered Emotion Stereotype

Brechet (2012) found that children stereotypically assign emotions to genders as well [5]. They assigned French children aged 6-8 years old, a completion task of gendered faces. The children had to draw the face of a character in a scenario sketched by the researchers. The scenarios were designed to elicit either anger or sadness in the characters. Boys tended to draw more angry faces, while girls tended to draw more sad or neutral faces. Both boys and girls were sensitive to the gender of the character, with angry male characters being drawn with a higher expressive intensity more often than the angry female characters that were drawn.

2.4 ChiSCor

The dataset containing the fantasy stories that are used in this thesis, ChiSCor was released by Van Dijk et al. (2023) [3]. Their way of collecting stories was based on the Story-Acting paradigm [6, 7]. They describe the ChiSCor corpus as a possible tool for studying the way children provide character perspectives throughout their cognitive development [8, 9]. This makes the corpus suitable for studying the differences in the way children describe male and female characters.

Van Dijk et al. (2023) have already done a case study to investigate gender bias in the ChiSCor corpus using a Word2Vec model [3]. It computed the difference in cosine similarity of the words ‘work’ or ‘home’ with ‘mommy’ vs. ‘daddy’. The differences were small which indicates a small gender bias in the ChiSCor corpus. This thesis further explores the existence of this gender bias, with a focus on the descriptions of the following categories: intelligence, emotions, personality, abilities, and physical appearance.

2.5 Definitions

The categories of descriptions that are studied in this thesis are intelligence, emotion, appearance, abilities, and personality. Table 1 shows the definitions of the categories according to the Cambridge Dictionary and examples of the words and phrases of descriptions accepted in this paper.

Category	Definition	Description examples
Intelligence	"The ability to learn, understand, and make judgments or have opinions that are based on reason." [10]	Smart, dumb, does not know, knows a lot, etc.
Emotion	"A strong feeling such as love or anger, or strong feelings in general." [11]	Happy, sad, angry, jealous, etc.
Appearance	"The way a person or thing looks to other people." [12]	Big, small, pretty, ugly, etc.
Ability	"The physical or mental power or skill needed to do something." [13]	Can do, can't do, good at, bad at, etc.
Personality	"The type of person you are, shown by the way you behave, feel, and think." [14]	Sweet, strict, honest, funny etc.

Table 1: Definitions according to Cambridge Dictionary and descriptions examples per category

3 Methods

This chapter describes the methodology of the current study. First, the sampling methods that were used to select stories from the ChiSCor Corpus are documented. Then an explanation and justification of the content analysis that is applied in this thesis. Furthermore, an overview of the annotation scheme used for the coding of the stories is given. Lastly, the statistical analysis methods are described.

3.1 Sampling methods

The ChiSCor corpus contains 619 fantasy stories told by 481 children aged 4-12. A detailed description of the sampling methods used to collect the stories can be found in Van Dijk et al. (2023) [3]. A brief summary is provided here; The data were collected in 2020-2022 at various primary schools, including one international school, a daycare, and a community center. The stories were collected within a natural context. The children were not influenced by pictures or feedback from interlocutors, this ensures that the stories they told are an expression of their own imagination and world views. The stories are also socially embedded since the children made up and told their fantasy stories in front of their classmates. Since the stories are told, the children have no opportunity to read back or change details. This natural context distinguishes the ChiSCor corpus from other datasets of children’s stories.

From the 619 stories within the ChiSCor corpus, one story from each participating child was selected. This decision is based on time constraints and on the fact that this method ensures equal representation for every child in the analysis. In the cases where a child has told multiple stories, the story with the largest file size is selected. This selection method aims to optimize the opportunity for increased character descriptions. If the biggest stories have the same file size, the story with the largest number of words is selected.

3.2 Content Analysis

Content analysis is used to quantify the number of character descriptions into categories, more information about the division of the categories can be found in the next section. The coding of the stories is done manually in this research. Manual coding is chosen due to the manageable sample size and the advantages of recognizing grammatical errors and changes in character names during the storytelling process. Since the children told these stories spontaneously, such inconsistencies occur often. Therefore, the decision was made to manually code the stories rather than coding them computationally. However, it is important to acknowledge the potential for human bias when interpreting the results, considering the manual coding approach.

3.3 Annotation Scheme

The manual coding was done with the help of an annotation scheme. This annotation scheme was developed using the insight gained from reading and attempting to code a selection of 15 stories. These stories were selected relatively randomly by scrolling through the database and randomly selecting stories. The annotation scheme was unaltered after this. All the stories used to develop the annotation scheme are coded again with the final version of the annotation scheme. The annotation

scheme can be found in Appendix .1. In Figure 1, a schematic overview of the annotation scheme is given. In addition to the categories shown in this overview, the annotation scheme also notes if the character is referred to in the first person and the occupations of the character.

The categories of descriptions as mentioned before are present in the annotation scheme. For each of these categories is a further division made into positive, negative, and neutral. This refers to the connotations a description has in the story. Consequently, big could be positive, negative, or neutral depending on the story.

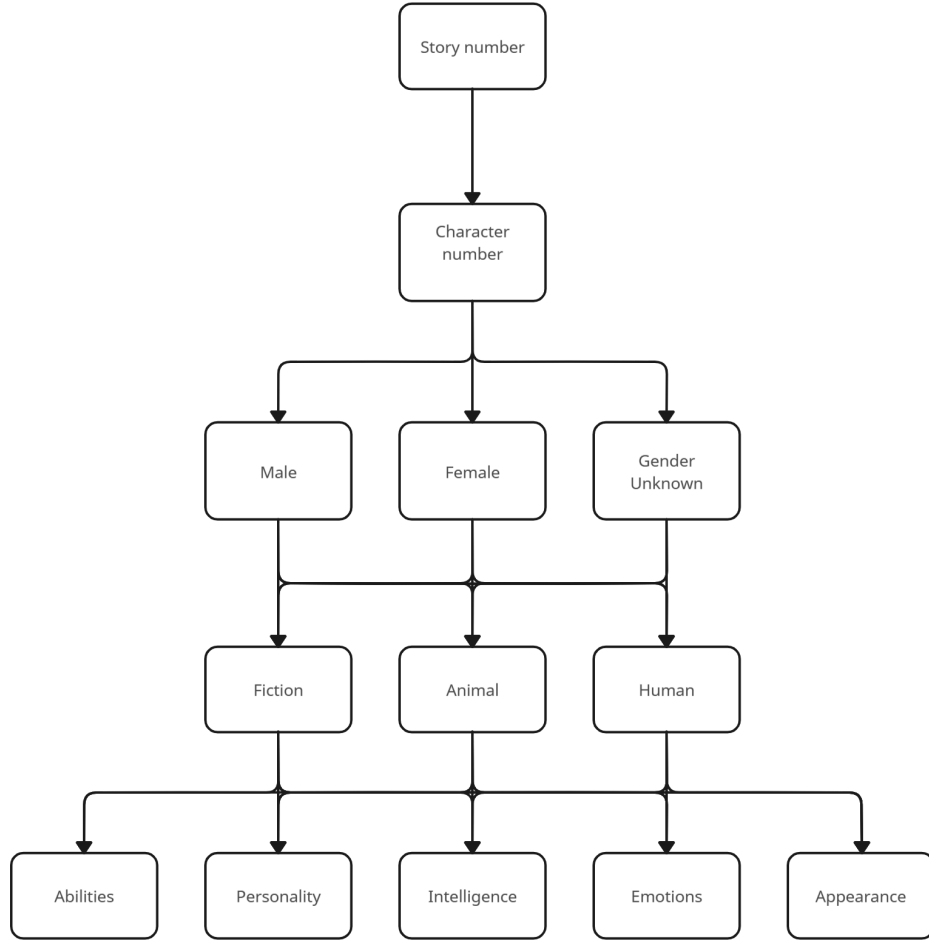


Figure 1: Schematic Overview of Annotation Scheme.

3.4 Statistical analysis

This section describes the statistical analysis used to study if there is a statistical difference between the number of male and female characters, between the number of descriptions of these male and female characters per category, and between the connotations of every category's descriptions in male and female characters.

3.4.1 Average number of female vs male characters per story

The following analysis aims to identify if there is a significant difference in the representation of male and female characters. The analysis consists of 2 two statistical tests. A Levene's Test for homogeneity of variances to validate the assumption of equal variances, which is a prerequisite for the following Independent Two-Sample T-Test. This t-test compares the average number of male and female characters across all stories.

3.4.2 Emotional Descriptions by Gender

The first phase of the statistical analysis assesses the association between the character's gender and the presence of emotional descriptions in the ChiSCor database. Proportions of characters with and without descriptions are calculated for each gender and illustrated with a contingency table. A Chi-squared test is then used to compare the proportions between male and female characters.

The second phase assesses the association between types of emotional descriptions (positive, negative) and character genders (female, male, and unknown). Proportions of positive and negative descriptions are calculated for each gender, based on the number of characters of the corresponding gender that received an emotional description. A Chi-Square Test is then performed on a contingency table comprised of the counts of each type of emotional description across the gender categories, to investigate any significant associations.

3.4.3 Intelligence Descriptions by Gender

The first phase of this statistical analysis assesses the association between the character's gender and the presence of intelligence descriptions in the ChiSCor database. Proportions of characters with and without descriptions are calculated for each gender and illustrated with a contingency table. A Chi-squared test is then used to compare the proportions between male and female characters.

The second phase involves examining the association between the type of intelligence descriptions (positive, negative) and character gender (female, male, unknown). Proportions of each description type are calculated for each gender category. Fisher's Exact Test is then conducted on a contingency table comprised of the counts of each type of emotional description across the gender categories, to investigate any significant associations. Due to the small number of descriptions, Fisher's Exact Test is chosen over the Chi-squared Test of Independence.

3.4.4 Appearance Descriptions by Gender

The first phase of this statistical analysis assesses the association between the character's gender and the presence of appearance descriptions in the ChiSCor database. Proportions of characters with and without descriptions are calculated for each gender and illustrated with a contingency table. A Chi-squared test is then used to compare the proportions between male and female characters.

The following analysis examines the association between the type of intelligence descriptions (positive, negative, neutral) and character gender (female, male, unknown). All three connotations, positive, negative, and neutral are included because of the large number of neutral appearance

descriptions. Proportions of each description type are calculated for each gender category. A Chi-squared Test of Independence is conducted to determine if there is a significant association between the gender of characters and these description types.

3.4.5 Ability Descriptions by Gender

The first phase of this statistical analysis assesses the association between the character's gender and the presence of ability descriptions in the ChiSCor database. Proportions of characters with and without descriptions are calculated for each gender and illustrated with a contingency table. A Chi-squared test is then used to compare the proportions between male and female characters.

The following analysis examines the association between the type of ability descriptions (positive, negative) and character gender (female, male, unknown). Proportions of positive and negative descriptions were computed for each gender category. Due to the small number of descriptions, Fisher's Exact Test was chosen over the Chi-squared Test of Independence. This test was applied to assess the association between gender and the type of ability descriptions.

3.4.6 Personality Descriptions by Gender

The first phase of this statistical analysis assesses the association between the character's gender and the presence of personality descriptions in the ChiSCor database. Proportions of characters with and without descriptions are calculated for each gender and illustrated with a contingency table. A Chi-squared test is then used to compare the proportions between male and female characters.

The following analysis examines the association between the type of personality descriptions (positive, negative) and character gender (female, male, unknown). Proportions of positive and negative descriptions are computed for each gender category. A Chi-Square Test is then performed on a contingency table comprised of the counts of each type of personality description across the gender categories, to investigate any significant associations.

3.4.7 Occupations

A superficial examination of the occupations assigned to the characters is done by looking at three word clouds. The first shows the occupations assigned to all characters. The second and third word clouds show the occupations assigned to female and male characters respectively.

4 Results

A total of 481 stories were coded. In these stories, 1536 characters were mentioned. 183 of these characters had one or more emotion descriptions, 32 characters had one or more intelligence descriptions, 168 characters had one or more appearance descriptions, 79 characters had one or more ability descriptions, and 75 characters had one or more personality descriptions. This chapter describes the associations between the presence of the descriptions and the character's gender, using the results of the statistical tests described in the previous chapter.

4.1 Average number of female vs male characters per story

Summary statistics indicated the following average counts per story: females - 0.684, males - 1.135, unknown gender - 1.374. A Levene's Test showed no significant difference in the variances between the number of male and female characters ($F(1, 960) = 1.278, p = 0.2586$). This result legitimized the subsequent independent two-sample t-test.

The independent two-sample t-test revealed a statistically significant difference in the average number of male and female characters per story ($t(960) = -6.0586, p < 0.001$). Specifically, the findings indicated that, on average, stories contained more male characters (mean = 1.135) than female characters (mean = 0.684).

4.2 Emotional Descriptions by Gender

This next part assesses the association between the character's gender and the presence of emotional descriptions in the ChiSCor database. In Table 2, a contingency table that categorizes characters into two groups per gender, those with at least one emotional description and those without, is shown. It indicates that female characters are more likely to have emotional descriptions compared to the male and unknown gender characters. A Pearson's Chi-squared test on this contingency table reveals a significant association between the character's gender and the presence of emotional descriptions ($X^2 = 18.52, df = 2, p\text{-value} = 9.515e-05$).

	Without Emotional Description	With Emotional Description	Total
Female	274	55	329
Male	471	75	546
Unknown	608	53	661
Total	1353	183	1536

Table 2: Distribution of Character Descriptions by Gender and Presence of Emotion Description

Figure 2 illustrates a noticeable difference in the proportions of emotional connotations used to describe female, male, and unknown-gender characters. Notably, male characters have a higher proportion of negative emotional descriptions compared to female characters. Characters of unknown gender exhibit a relatively equal proportion of positive and negative emotional descriptions.

This observed distribution is supported by the results of the Pearson Chi-squared test ($X^2 =$

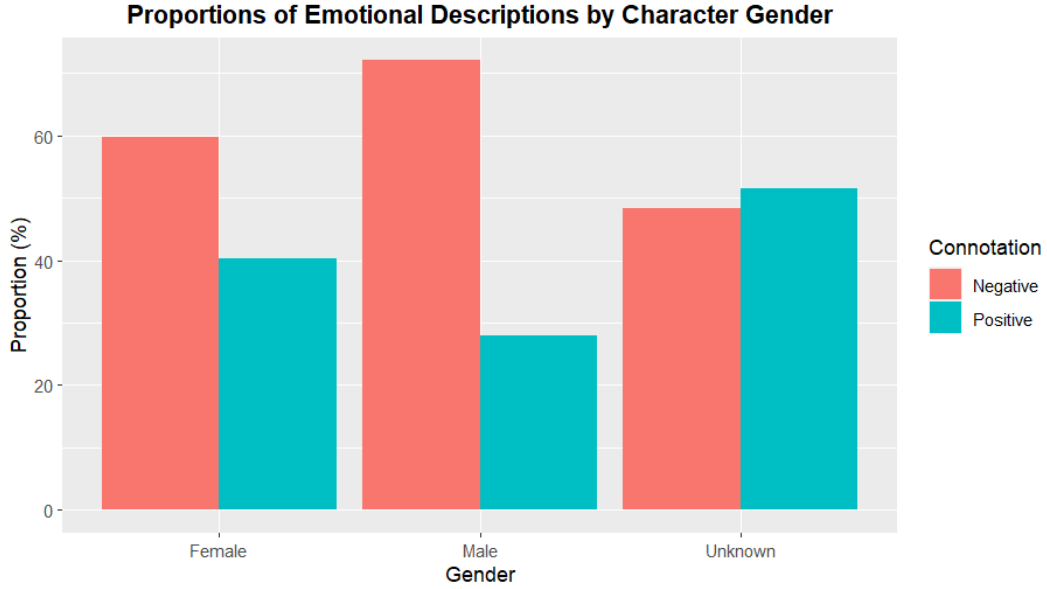


Figure 2: Connotation Proportions of Emotional Descriptions by Character Gender.

9.5771, $df = 2$, $p\text{-value} = 0.008325$), which indicates a statistically significant association between the gender of the characters and the type of emotional descriptions they receive.

4.3 Intelligence Descriptions by Gender

This next part assesses the association between the character’s gender and the presence of intelligence descriptions in the stories. In Table 3, a contingency table that categorizes characters into two groups per gender, those with at least one intelligence description and those without, is shown. It indicates that female characters are more likely to have intelligence descriptions compared to the male and unknown gender characters. A Pearson’s Chi-squared test on this contingency table reveals a significant association between the character’s gender and the presence of intelligence descriptions ($X\text{-squared} = 8.1104$, $df = 2$, $p\text{-value} = 0.01733$).

	Without Intelligence Description	With Intelligence Description	Total
Female	316	13	329
Male	535	11	546
Unknown	8	8	661
Total	1504	32	1536

Table 3: Distribution of Character Descriptions by Gender and Presence of Intelligence Description

Figure 3 displays the proportions of positive and negative intelligence descriptions by gender. It shows that male characters are more frequently associated with negative intelligence descriptors

and that female characters are more commonly linked with positive ones. The unknown gender category was also more frequently described with positive descriptions.

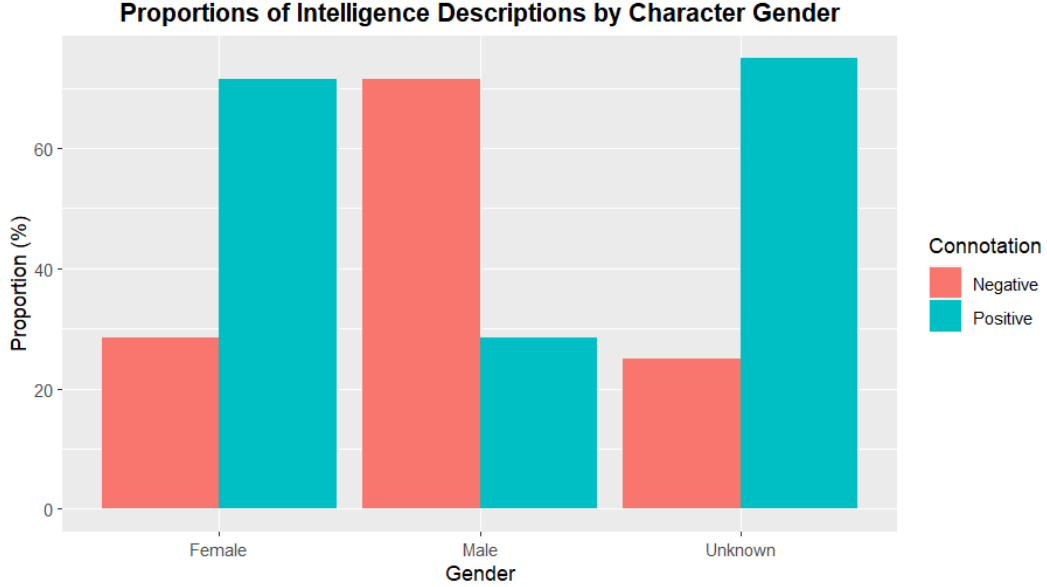


Figure 3: Connotation Proportions of Intelligence Descriptions by Character Gender.

Fisher’s Exact Test ($p\text{-value} = 0.03945$) demonstrated a significant association between character gender and the type of intelligence descriptions, underlining a gender-based inequality in the description of intelligence.

4.4 Appearance Descriptions by Gender

This next part assesses the association between the character’s gender and the presence of appearance descriptions in the ChiSCor database. In Table 4, a contingency table that categorizes characters into two groups per gender, those with at least one appearance description and those without, is shown. It indicates that male characters are more likely to have appearance descriptions compared to female and unknown-gender characters. A Pearson’s Chi-squared test on this contingency table reveals a significant association between the character’s gender and the presence of appearance descriptions ($X\text{-squared} = 8.7144$, $df = 2$, $p\text{-value} = 0.01281$).

	Without Appearance Description	With Appearance Description	Total
Female	299	30	329
Male	469	77	546
Unknown	600	61	661
Total	1368	168	1536

Table 4: Distribution of Character Descriptions by Gender and Presence of Appearance Description

Figure 4 shows the proportions of positive, negative, and neutral appearance descriptions by gender. It shows that female and unknown-gender characters are most commonly described with neutral descriptions. Male characters are more frequently described with negative appearance descriptions, followed relatively closely by the frequency of neutral descriptions.

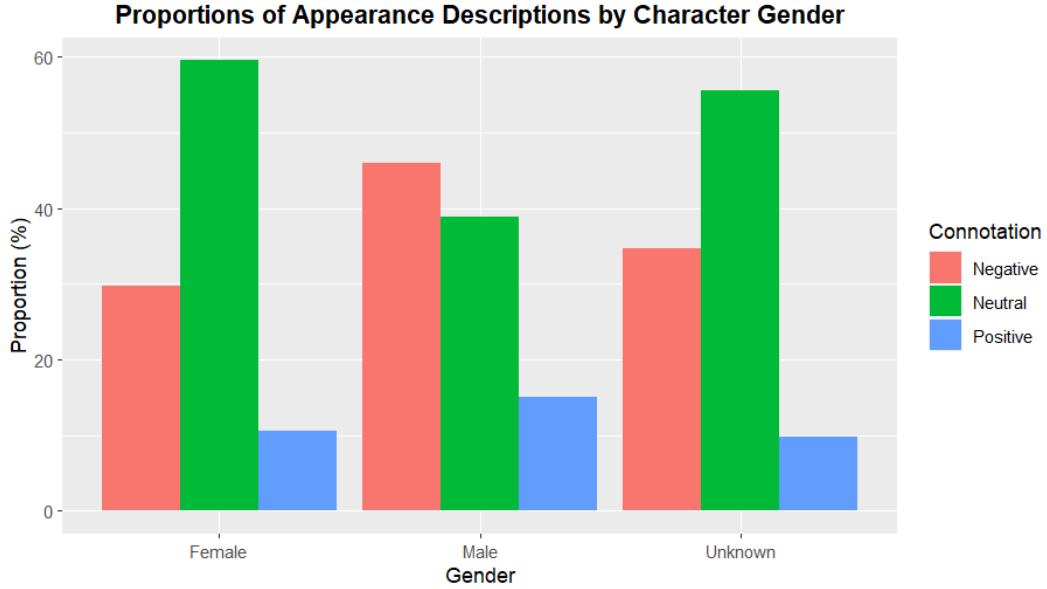


Figure 4: Connotation Proportions of Appearance Descriptions by Character Gender.

The Chi-squared test of independence was then conducted to test if the association between character gender and the connotations of physical appearance descriptions is statistically significant ($X^2 = 8.4844$, $df = 4$, $p\text{-value} = 0.07536$). Although this p -value did not reach the conventional threshold for statistical significance ($p < 0.05$), it suggests a marginal trend that warrants further investigation.

4.5 Ability Descriptions by Gender

This next part assesses the association between the character's gender and the presence of ability descriptions in the ChiSCor database. In Table 5, a contingency table that categorizes characters into two groups per gender, those with at least one ability description and those without, is shown. It indicates that male characters are a little bit more likely to have ability descriptions compared to female and unknown-gender characters. A Pearson's Chi-squared test on this contingency table does not reveal a significant association between the character's gender and the presence of ability descriptions ($X^2 = 5.7424$, $df = 2$, $p\text{-value} = 0.05663$).

Figure 5 shows a relatively similar distribution of positive and negative ability descriptions among all gender categories. With the proportion of positive ability descriptions being bigger for every gender. The difference in negative and positive description proportions is the smallest for female characters, followed by male characters and unknown gender characters have the biggest negative proportion.

To test if the association between character gender and the connotations of ability description

	Without Ability Description	With Ability Description	Total
Female	315	14	329
Male	508	38	546
Unknown	634	27	661
Total	1457	79	1536

Table 5: Distribution of Character Descriptions by Gender and Presence of Ability Description

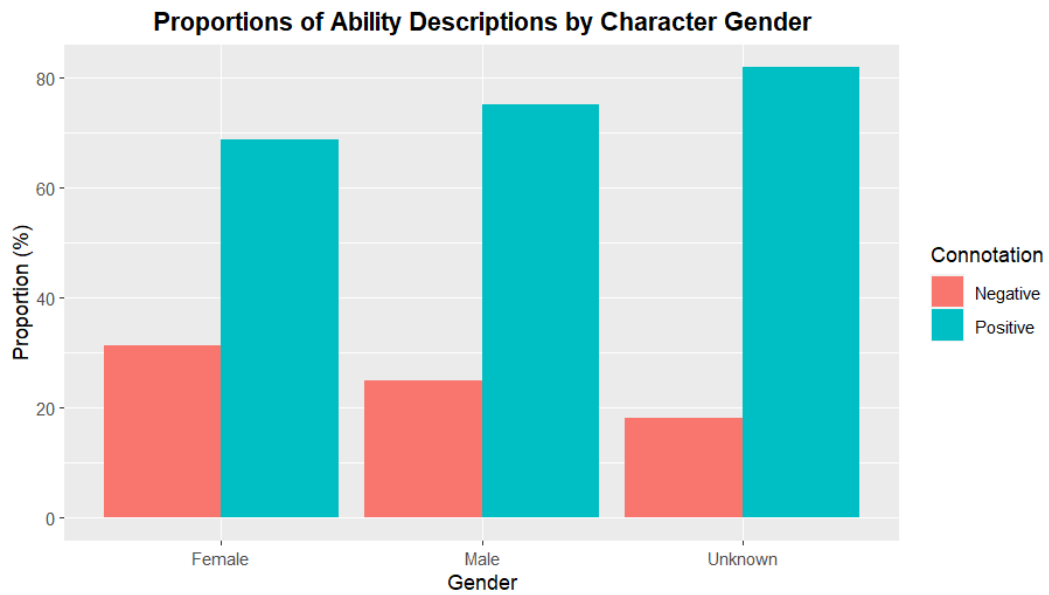


Figure 5: Connotation Proportions of Ability Descriptions by Character Gender.

is statistically significant, a Fisher’s Exact Test was executed. The resulting p-value (p-value = 0.5902). suggests no statistically significant association between gender and the type of ability descriptions.

4.6 Personality Description by Gender

This next part assesses the association between the character’s gender and the presence of personality descriptions in the ChiSCor database. In Table 6, a contingency table that categorizes characters into two groups per gender, those with at least one personality description and those without, is shown. It indicates that female characters are more likely to have personality descriptions than male and unknown-gender characters. Unknown gender characters are least likely to have a personality description compared to both female and male characters. A Pearson’s Chi-squared test on this contingency table reveals a significant association between the character’s gender and the presence of personality descriptions (X-squared = 8.7024, df = 2, p-value = 0.01289).

	Without Personality Description	With Personality Description	Total
Female	305	24	329
Male	516	30	546
Unknown	640	21	661
Total	1461	75	1536

Table 6: Distribution of Character Descriptions by Gender and Presence of Personality Description

Figure 6 illustrates the distribution of positive and negative personality traits across genders. It shows a tendency for male characters to be associated with positive personality traits more than female characters, while unknown-gender characters have the highest proportion of positive descriptions.

However, despite what the figure visually suggests about male characters having a higher proportion of negative descriptions, the Chi-squared test did not find this difference to be statistically significant (p-value = 0.4087). This indicates that the observed difference in the proportions is not strong enough to be statistically proven.

4.7 Occupations

A superficial examination of the occupations assigned to each gender can be done using the word clouds presented below. Figure 7 displays a word cloud of the occupations assigned to characters, while Figures 8 and 9 further divide the assigned occupations for female and male characters, respectively.

As seen in the word cloud in Figure 7, the occupation assigned most frequently is ‘moeder’ (Mother), with a total of 45 times, followed by ‘ridder’ (knight), which is used 37 times. 26 words were used to assign occupations to female characters. From the occupations assigned to female characters ‘moeder’ (mother) is most frequently used followed by ‘mama’ (mommy), 28 times, and ‘prinses’ (princess),

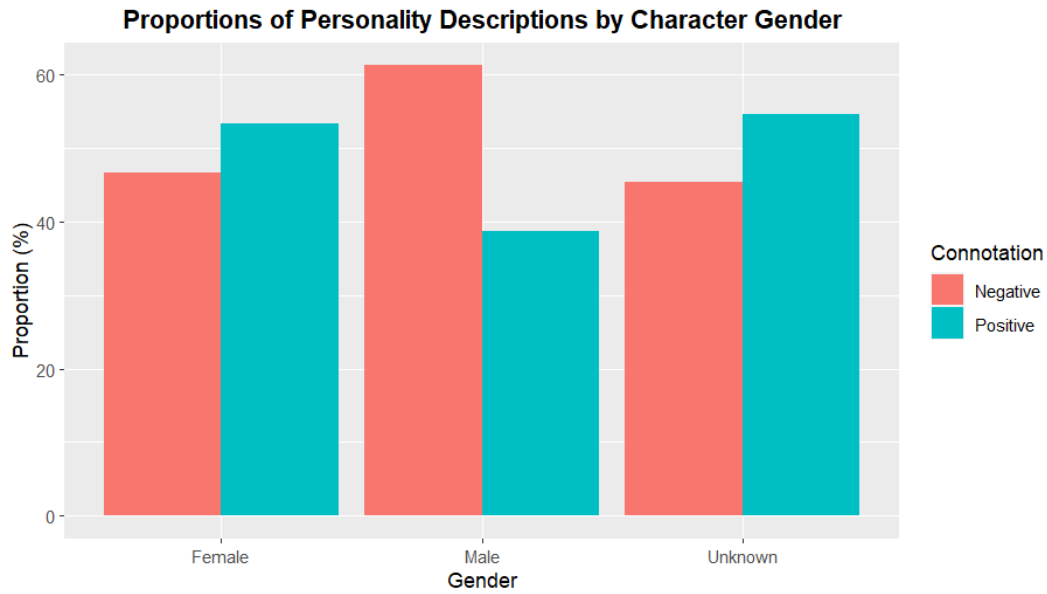


Figure 6: Connotation Proportions of Personality Descriptions by Character Gender.

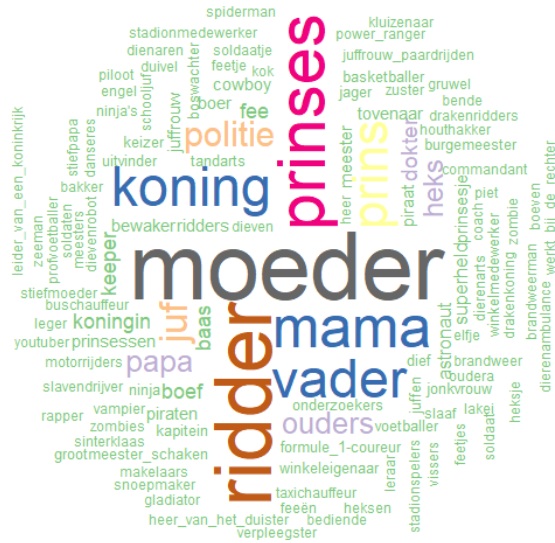


Figure 7: Occupations Assigned to all Characters.

Characters

Characters

26 times. 56 words were used to assign occupations to male characters. From the occupations assigned to male characters, 'vader' (father) was assigned most frequently, 23 times, followed closely by 'koning' (king), 22 times, and 'prins' (prince), 20 times.

5 Discussion

This study sets out to explore the differences in the way different genders are described by children using stories from the ChiSCor database. The findings indicate significant differences in how often emotions, intelligence, physical appearance, and personalities are described for characters of different genders. No significant associations between gender and ability descriptions are found in this thesis.

The stories contain on average more male characters than female characters. This finding points out a gender imbalance in character representation in stories told by children. This is similar to the imbalance that exists in children’s books.

The study also reveals significant associations between character gender and the presence of at least one emotional description, with females being more likely to have one or more emotional descriptions. This confirms the hypothesis that the stereotype that females are more emotional is present in the ChiSCor database. There was also a significant association between gender and the type of emotional description, with male characters having a higher proportion of negative emotions than female characters. Both male and female characters have a higher proportion of negative descriptions than positive descriptions.

For the personality descriptions, there is a significant association between character gender and the presence of at least one personality description, with females being more likely to have one or more personality descriptions. No significant association between gender and the type of personality description is found.

The study reveals significant associations between character gender and the presence of at least one appearance description and significant associations between character gender and the type of appearance description (positive, negative, or neutral). Male characters are more likely to have an appearance description than female characters. They are also more commonly linked with negative appearance descriptions closely followed by neutral appearance descriptions. Females are most commonly described with a neutral description.

A statistically significant association between character gender and the presence of at least one intelligence description is found, with females being more likely to have one or more intelligence descriptions than male characters. There is also a significant association between gender and the type of intelligence description. Males are more commonly linked with negative descriptions than positive descriptions. And females have a higher proportion of positive intelligence descriptions than negative descriptions. Which is contrary to the hypothesis of the “brilliance = male” being present in the corpus. The results indicate an opposite stereotype in the data. It could also be that male intelligence is perceived as implicit and the children did not feel the need to point it out.

The superficial examination of the occupations assigned to characters shows that parents are mentioned the most. This is intuitively right since parents are a big part of children’s worlds. It also showed that there is a bigger variety of occupations assigned to male characters compared to female characters. However, there are more male characters than female characters mentioned in the stories. Future research should examine this difference further to see if there is a significant

association between the variety of occupations and character gender.

This thesis acknowledges certain limitations that may impact the interpretation of the findings:

- The influence of the narrator's gender and age on variations in character descriptions was not considered in this study. Future research could explore these factors to provide a more comprehensive understanding.
- The reliance on a single coder introduces the possibility of subjective interpretation. Future studies could involve multiple coders to enhance the robustness and reliability of the analysis.
- The study did not investigate how the gender and age of narrators might influence character descriptions. Exploring these demographic variables could provide valuable insights into potential biases in storytelling.
- The influence of character types (human, animal, fictional) on descriptions was not examined. Future research should delve into whether these distinctions impact how characters are portrayed by children.
- Although the literature review provided a foundation for the study, a more extensive exploration of existing literature could improve the contextual understanding and strengthen the link between this research and prior research.
- The study did not analyze the impact of the number of words spoken by characters on the portrayal of gender stereotypes. Future investigations could explore this aspect to gain a more nuanced perspective.

Continuing to address these aspects in future research will contribute to a more complete understanding of how and to what extent gender stereotypes are internalized in children.

6 Conclusion

In conclusion, this study explores the differences in the way different genders are described by children using stories from the ChiSCor database. The research reveals significant differences in how emotions, intelligence, physical appearance, and personalities are described for characters of different genders. The findings shed light on the presence of gender stereotypes in children's narratives. In acknowledging the persistent gender inequalities in the Netherlands, this study emphasizes the importance of early interventions in challenging stereotypes. By understanding the gender stereotypes internalized by children, educators, and policymakers can work towards breaking harmful gender stereotypes and promoting diversity in future generations.

The findings and limitations presented in this thesis serve as a foundation for future studies, urging researchers to continue examining the gender stereotypes in children's narratives, to be able to foster an environment where every child can envision themselves beyond traditional roles and expectations.

References

- [1] L. Bian, S.-J. Leslie, and A. Cimpian, “Gender stereotypes about intellectual ability emerge early and influence children’s interests,” *Science*, vol. 355, no. 6323, pp. 389–391, 2017.
- [2] “She figures 2021.”
- [3] B. M. A. van Dijk, M. J. van Duijn, S. Verberne, and M. R. Spruit, “Chiscor: A corpus of freely told fantasy stories by dutch children for computational linguistics and cognitive science,” 2023.
- [4] T. L. King, A. J. Scovelle, A. Meehl, A. J. Milner, and N. Priest, “Gender stereotypes and biases in early childhood: A systematic review,” *Australasian Journal of Early Childhood*, vol. 46, no. 2, pp. 112–125, 2021.
- [5] C. Brechet, “Children’s gender stereotypes through drawings of emotional faces: Do boys draw angrier faces than girls?,” *Sex Roles*, vol. 68, pp. 378–389, Mar 2013.
- [6] T. Cremin, R. Flewitt, B. Mardell, and J. Swam, *Storytelling in Early Childhood: Enriching language, literacy and classroom culture*. London: Routledge, 2016.
- [7] A. Nicolopoulou, K. S. Cortina, H. Ilgaz, C. B. Cates, and A. B. de Sá, “Using a narrative-and play-based activity to promote low-income preschoolers’ oral language, emergent literacy, and social competence,” *Early childhood research quarterly*, vol. 31, pp. 147–162, 2015.
- [8] A. Nicolopoulou and B. Ünlütürk, “Narrativity and Mindreading Revisited: Children’s Understanding of Theory of Mind in a Storybook and in Standard False Belief Tasks,” in *Social Environment and Cognition in Language Development*, pp. 151–166, John Benjamins, 2017.
- [9] B. Van Dijk and M. Van Duijn, “Modelling Characters’ Mental Depth in Stories Told by Children Aged 4-10,” in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 43, pp. 2384–2390, 2021.
- [10] *Cambridge Dictionary - intelligence*. 1 2024.
- [11] *Cambridge Dictionary - emotion*. 1 2024.
- [12] *Cambridge Dictionary - appearance*. 1 2024.
- [13] *Cambridge Dictionary - ability*. 1 2024.
- [14] *Cambridge Dictionary - personality*. 1 2024.

.1 Annotation Scheme

See ‘annotation scheme.csv’ for the annotations made on ChiSCor, which is publicly available via the Open Science Framework link as provided in Van Dijk et. al., 2023 [\[3\]](#).