

Does Shakespeare Pass the Bechdel Test?

Zhiyan Foo

November 11, 2016

Abstract

The Bechdel Test is measure of gender bias in a piece of artistic work. This paper presents the findings when the Bechdel Test is applied to Shakespeare’s plays computationally. The results of the experiment show that Shakespeare’s plays are overwhelmingly focused on the depiction of the lives of its male characters.

1 Background and Significance

Despite the recent success of films such as *The Hunger Games* and *Star Wars: The Force Awakens*, most of our cultural output is disproportionately focused on the male lives and their stories.[2]. One popular measure of this gender bias is the Bechdel Test[3]. The test specifies 3 criteria for a piece to pass.

1. It must contain two named female characters
2. They must have a conversation
3. The conversation must not be about a man

While this is not a particularly high bar, a surprising amount of films—the original medium the test was applied to¹—fail it. Gender disparities however are not a recent phenomena and this article looks at whether Shakespeare’s plays in particular pass the Bechdel Test. This would not be done manually however but as a showcase for a new software package, **crunch-shake**².

¹see <http://bechdeltest.com/> for a crowdsourced effort to classify movies by the Bechdel Test.

²<https://github.com/zhiyanfoo/crunch-shake>

2 Methodology

2.1 Modified Criteria

The criteria used by crunch-shake is a modified version of the original Bechdel Test. This is to both accommodate and utilize the computational nature of crunch-shake.

Original Criteria	crunch-shake Criteria	Reason for Change
Contain two named female characters	Contain two female characters that are in the upper 50% of notable characters. See Gender and Notability .	The requirement that the two female characters be named is just a proxy for whether the character is significant to the piece, and using social network, graphing algorithms and the number of lines the character has, significance of the character can be determined directly.
The two named female characters must have a conversation	In any scene, two notable female characters must speak in the presence of one another. See Presence .	It is hard to algorithmically determine when two characters are in a 'conversation' with one another. So the two females might be talking to a male, not each other, but unfortunately the scene will still 'pass'.
The conversation must not be about a man	In their 'conversation', the two notable females must not utter a word related to romantic relationships or mention a male. See Blacklist .	While there's a lot of subtext that an algorithm can miss out, a blacklist of words takes care of the more obvious cases.

Whereas traditionally, the Bechdel Test has only been applied to females, the computational nature of the test means that the test can be also easily be applied to males, and the results of the two tests compared. So for example the male version of the Bechdel Test would check to see if there exists at least two notable males who talk about something other than women in a scene.

2.2 Gender

The gender of each character is determined manually.³ While it is easy enough to determine the gender of the named characters, the unnamed characters proved to be a bit harder. Some gender assignments were easy, such as ‘groom’ or ‘maid’. Others such as ‘Soldier’, while could represent both genders in today’s society, were for the most part male in Shakespeare’s time and so classified as such. The more ambiguous designations such as ‘Citizen’ while were for the most part were probably intended to be male, are left as undetermined, and for the purposes of the algorithm, these people might as well not exist.

2.3 Notability

Whether a character is notable or not is dependent on how each character scores on 4 metrics: lines by character, out degree, page rank, betweenness. The last 3 metrics are network algorithms. crunch-shake uses the implementation found in the python package networkx⁴.

In this case, the network vertices here are speaking characters, with directed edges representing when a character speaks to another character, with more lines of speech indicating a stronger connection.

The use of network algorithms to classify importance of characters was taken from a paper on Game Of Thrones(GOT)[1], a TV series. Since Shakespeare’s plays feature far fewer characters than in GOT and since those characters are far more interconnected than in GOT, this methodology does not work as well for this paper, which is why the network algorithms are assigned such low weights, as shown in Table 1.

³<https://github.com/zhiyanfoo/crunch-shake/tree/master/crunch-shake/gender>

⁴<https://networkx.readthedocs.io/en/stable/reference/algorithms/centrality.html>

Table 1: Metrics used by crunch-shake to evaluate notability

Metric	Weight	Algorithm
lines by character	62.5%	Take the number of lines a character has and divide it by the lines of speech the character with the most lines has.
out degree	12.5%	networkx’s implementation
page rank	12.5%	networkx’s implementation
betweenness	12.5%	networkx’s implementation

2.4 Presence

2.4.1 Methods

There are two ways the algorithm in crunch-shake knows that a character is in a scene. The first is through stage directions. For example, if the algorithm sees the stage direction,

Enter CAPULET in his gown, and LADY CAPULET.⁵

it would note that CAPULET and LADY CAPULET have entered. Similarly, the algorithm will remove the character from the scene if it sees “Romeo Exit”

In the second method, if the algorithm sees that Romeo speaks a line, even if he never ‘entered’ the scene—so that scene started *in media res*—it would take it that Romeo was always there. Romeo would be assumed to be present until it it sees “Romeo Exit” or the scene ends.

2.4.2 Limitations

There are limitations to the algorithm however. For example if it sees,

Exeunt all but MONTAGUE, LADY MONTAGUE, and BEN-VOLIO.⁶

⁵From Act I, Scene I in *Romeo and Juliet*

⁶From the same scene

it would erroneously note that MONTAGUE, LADY MONTAGUE, and BENVOLIO have all exited as it sees 'exeunt' while keeping the rest of the characters as it is not smart enough to interpret 'but all'.

Also if a character enters, exit and re-enters, the algorithm will note only the first entrance and the last exit. Finally sometimes the play directions do not refer to the characters by name. Take this example from Act IV, Scene III of *The Taming of the Shrew*, where the SERVANT is not mentioned directly,

Enter four or five Serving-men

yet is suppose to enter the scene. These errors are unlikely to skew the data in favor of males or females however.

2.5 Blacklist

The list of forbidden words is tabulated below.

romantic relationships marriage, matrimony, courting, love, wedlock, sex, sexual, intercourse

partners partner, spouse, lover, admirer, fianc, amour, inamorato, betrothed.

specific to females boyfriend, husband

specific to males girlfriend, wife

In addition the names of all the male or female characters would also be blacklisted. The idea is not that it is 'wrong' for script writers to ever have their characters mention the opposite sex, but rather not ever scene involving two people from the same gender should have them discussing the opposite gender.

2.6 Precaution

In order not to bias the data by, for example changing the blacklisted words, or changing the weights given to each metrics, I wrote the specifications before running crunch-shake on Shakespeare's plays, with the exception of *All's Well that Ends Well* and *Romeo and Juliet*, which I needed for debugging and checking feasibility purposes. As a result I will be omitting these two plays from the results.

3 Results

3.1 Evaluation of Results

Even though I wasn't expecting Shakespeare to give equal weight to both sexes in his plays, the extreme discrepancies that crunch-shake found still surprised me ⁷ For females, 27 out of 35 of his plays do not pass the Bechdel Test ⁸, whereas for males, only one of the plays, *Love's Labour's Lost*, fails the Bechdel Test. And even that was only because the men in it managed to talk about women in every scene. However it needs to be acknowledged the failure rate of the Bechdel Test for males per scene—as opposed to per play—was quite high. In particular, 67% of the scenes did not pass the Bechdel Test for males⁹. Not nearly as bad as the figure for females, 98% , but not a small figure either. However if you take a closer look at the make up of these figures, it is evident that the scenes failed for females and males for different reasons. In 47% of the scenes, males failed because they talked about women¹⁰, whereas the figure is 17% for women. This is not because women tend not to talk about men, in fact in 91% of the scenes in which they talk, they mention men. This figure is 58% for men. However in general women don't get to talk to each other very often. In 80% of the scenes, there was no conversation between two notable females ¹¹, whereas this occurred on 20% of the time for males.

⁷For the figures in Table 3 (SFB, SNC, SBL and COG) these are calculated by considering all scenes in Shakespeare's plays again, not by averaging the results in Table 2. This is so as to not diminish the impact of longer works while exaggerating the effect of shorter works.

⁸For a play to fail, all scenes must fail the Bechdel Test.

⁹If a scene failed, either the scene had 'no conversation', or the conversation included 'blacklisted' words

¹⁰crunch-shake considers a scene to have failed if if two members of the same gender have a conversation which mentions the opposite gender. It does this by having a blacklist of words the conversation must not include.

¹¹crunch-shake determines a conversation to have occurred if two notable characters from the same gender both speak within a scene.

Table 2: Ranking of Shakespeare’s Play’s by FB (F) %

Rank	Play	FB (F) %	FB (M) %	NC (F) %	NC (M) %	BL (F) %	BL (M) %
1	Comedy of Errors	90.91	81.82	45.45	18.18	45.45	63.64
2	King Lear	92.31	61.54	80.77	15.38	11.54	46.15
3	Winter’s Tale	92.86	92.86	71.43	21.43	21.43	71.43
4	Henry IV, part 2	94.74	47.37	89.47	15.79	5.26	31.58
5	As You Like It	95.45	81.82	54.55	36.36	40.91	45.45
6	Henry V	95.65	43.48	91.30	13.04	4.35	30.43
7	Cymbeline	96.15	80.77	88.46	30.77	7.69	50.00
8	Antony and Cleopatra	97.62	57.14	76.19	23.81	21.43	33.33
9	Much Ado About Nothing	100.0	88.24	64.71	23.53	35.29	64.71
10	Macbeth	100.0	67.86	78.57	39.29	21.43	28.57
11	Timon of Athens	100.0	41.18	100.0	23.53	0.00	17.65
12	Twelfth Night	100.0	72.22	77.78	38.89	22.22	33.33
13	Henry IV, part 1	100.0	47.37	100.0	15.79	0.00	31.58
14	Richard III	100.0	56.00	76.00	28.00	24.00	28.00
15	Othello	100.0	86.67	60.00	6.67	40.00	80.00
16	Henry VI, part 3	100.0	39.29	100.0	7.14	0.00	32.14
17	Coriolanus	100.0	65.52	82.76	24.14	17.24	41.38
18	Midsummer Night’s Dream	100.0	77.78	44.44	0.00	55.56	77.78
19	The Tempest	100.0	66.67	66.67	0.00	33.33	66.67
20	Two Gentlemen of Verona	100.0	95.00	85.00	45.00	15.00	50.00
21	Merry Wives of Windsor	100.0	91.30	65.22	21.74	34.78	69.57
22	Measure for Measure	100.0	88.24	70.59	35.29	29.41	52.94
23	Hamlet	100.0	70.00	85.00	15.00	15.00	55.00
24	Henry VI, part 1	100.0	40.74	96.30	11.11	3.70	29.63
25	Julius Caesar	100.0	44.44	100.0	11.11	0.00	33.33
26	Troiles and Cressida	100.0	66.67	100.0	8.33	0.00	58.33
27	Richard II	100.0	57.89	100.0	21.05	0.00	36.84
28	Henry VI, part 2	100.0	41.67	100.0	4.17	0.00	37.50
29	Love’s Labour’s Lost	100.0	100.0	66.67	0.00	33.33	100.0
30	Taming of the Shrew	100.0	92.86	64.29	0.00	35.71	92.86
31	Henry VIII	100.0	88.24	94.12	17.65	5.88	70.59
32	Merchant of Venice	100.0	78.95	57.89	36.84	42.11	42.11
33	Titus Andronicus	100.0	71.43	92.86	0.00	7.14	71.43
34	King John	100.0	37.50	87.50	6.25	12.50	31.25
35	Pericles	100.0	84.21	73.68	15.79	26.32	68.42

Legend	FB	% Scenes that fail the Bechdel Test
	NC	% Scenes with 'No Conversations'
	BL	% Scenes with 'Blacklisted' conversation

Table 3: Results Across All Plays

	PFB	SFB %	SNC %	SBL %	COG %
Female	27	98	80	17	91
Male	1	67	58	20	58

Legend	PFB	Plays that fail the Bechdel Test
	SFB	Scenes that fail the Bechdel Test
	SNC	Scenes with ‘No Conversations’
	SBL	Scenes with ‘Blacklisted’ conversation
	COG	Conversation mentions the other gender

3.2 Examination of Scenes

Taking a closer look at the small number scenes which the passed for females, one realizes it might be too soon to even celebrate even those small achievements. Randomly choosing 5 out of the 8 scenes that passed, 2 of them (*Comedy of Errors* and *As You Like It*) do contain two women speaking, but it is about a man, they just don’t mention him by name. In *Henry V*, Katherine and her lady-in-waiting, Alice, do speak, but it is in French, which the algorithm does not understand. In *Winter’s Tale*, two women do speak, just not to each other but rather to the King of Sicily. Finally in *Henry IV, part 2*, a prostitute, Doll Tearsheet, and a brothel owner, Mistress Quickly, are not having a conversation as so much being dragged around by a beadle. As such none of these scenes can be said to have passed the Bechdel Test.

Looking at the scenes that passed for males, the first from *Henry VIII*, involves two noblemen making fun of the newest fashion from France. The second from *Henry IV, part 2*, involves Cade and the King’s ambassadors alternatively giving speeches to try get the mob to come to their respective sides. The third scene from *Othello* involves the eponymous protagonist giving instructions to his subordinates, as he is about to scout the camp’s fortification. The fourth scene, from *Pericles*, revolves around Thaliard eavesdropping on Helicanus, and subsequently lying to him. For the final scene, from *Henry VI, part 1*, the main exchange is actually between the Countess of Auvergne and Tabolt, and so can be constituted as a false positive.

3.3 Conclusion

Despite the already bleak figures generated by crunch-shake, indicating a huge disparity between the treatment of men and women by Shakespeare, the results could be even worse. A random sample of the scenes with females that crunch-shake passed were all shown to be false positives on closer examination—They all revolved around men in some way. In contrast an examination of the scenes with males, revealed such widespread scenes as two nobleman making fun of the newest french fashion to a general dispatching orders to his subordinates. Only one false positive was found in the five scenes examined.

Shakespeare is an eminent figure in English Literature, one could even say the eminent figure. While the themes he studies in his plays are universal, it is also clear that his plays are also singularly focused on the lives of men. While there are notable fleshed out females in Shakespeare’s play—Lady Macbeth and Portio from the *Merchant of Venice* come to mind—their values as characters are very much defined in relation to the more important male characters that dominate his plays. crunch-shake demonstrates not just on an anecdotal level, but on a statistical level, just how widespread this dominance is.

References

- [1] A. BEVERIDGE AND J. SHAN, *Network of thrones*, The American Mathematical Monthly.
- [2] S. JIA, T. LANSDALL-WELFARE, S. SUDHAHAR, C. CARTER, AND N. CRISTIANINI, *Women are seen more than heard in online newspapers*, PLoS ONE, (2016).
- [3] I. B. LYLE FRIEDMAN, MATT DANIELS, *Hollywood’s gender divide and its effect on films*, Polygraph.