

# Social Identity Theory and Voting Behaviour: FIFA Ballon d'Or 2015

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Department of Economics

Supervisor: Dana Sisak

Name: Mabshar Choudhury

Exam number: 364541

E-mail address: [mabshar\\_choudhury@hotmail.com](mailto:mabshar_choudhury@hotmail.com)

## Introduction

The FIFA Ballon d'Or award is a perfect event for a study in economic and psychology theories, the data is publicly available and easy to understand. The FIFA Ballon d'Or is an annual football award which is awarded to the world's best player by the FIFA (international governing body of football) and the French publication *France Football*. The FIFA Ballon d'Or is a fusion of the former Ballon d'Or and the FIFA World Player of the Year award. It is awarded based on votes from journalists, national team coaches and captains (FIFA, 2015). The award is widely seen as the most prestigious individual award in football. The aim of this project is to analyze data from the FIFA Ballon d'Or 2015 to determine whether voters base their choices on performances on the pitch or whether there some other factors that also play a role in their decision making process.

After every FIFA Ballon d'Or there a people who question the results, some raise questions about the voting choices, while others question the results because of FIFA's history with corruption. The purpose of the research is to determine whether the voting data reveals any biases and if so how the voting system can be improved to make it less prone to biases. Findings will be discussed in terms of the Social Identity Theory (SIT). This theory was developed by Henri Tajfel and John Turner in the 1970's. According to Tajfel and Turner people tend to favor members of the group they belong to over other groups (Tajfel & Turner, 1979) . In a perfect world the expectation would be that voters only would base their decisions on performances on the field of the candidates and no other factors.

The main research question for this thesis is: "Does the voting data for the FIFA Ballon d'Or 2015 award reveal any biases related to Social Identity Theory?"

Real life examples of voting's reveal that group forming leads to SIT related biases in voting's (Ben-Bassat & Dahan, 2012). SIT leads to individuals acting as a member of a group they identify themselves with instead of as an individual. These groups can be based on different factors, such as nationality, age or football team. In this thesis the focus lies on the difference in voting behavior between voters from different regions and of different ages.

In the first chapter of the thesis the SIT will be discussed in more detail and a overview of related literature is given. In the following chapter the research method en data used will be discussed. In chapter 4 the results are discussed which will be followed by a conclusion and discussion.

## Literature:

### **1.1 Social Identity Theory**

The social identity theory (SIT) was developed by Henri Tajfel and John Turner in the 1970's. Tajfel and Turner defined social identity as 'an individual's knowledge that he belongs to certain social groups and that these have some emotional and value significance to him. Groups can be defined as collections of people who share the same social identity. This leads to individuals acting as a member of a group they identify themselves with instead of as an individual. This means that for example when someone who identifies with a certain political party steps into the voting booth, they vote down the party line solely because they identify with the values and beliefs of that particular party (Yocco, 2014). According to this theory, the status of groups to which people believe they belong to helps them achieve a positive self-identity which creates in-group favoritism at the expense of the out-group (Sanderson, 2010). Research shows that in-group favoritism is stronger for groups that are formed naturally (e.g. race, sex) than for experimentally created groups (Doise, Csepele, Dann, Couge, Larsen, & Ostell, 1972). The desire to maintain a positive in-group status is higher for groups for which one has a higher degree of identification (i.e., race/ethnicity), which means that people are more motivated to protect the status and interests of these kind of groups (McKinley, Mastro, & Warber, 2014).

### **1.2 Social Identity Theory in voting**

The role of SIT in decision making has been the subject quite some research over time. Several researchers have found evidence that SIT plays a role in decisions making which leads to biases in the voting results. One of the most famous real life examples of SIT in voting comes from a research done by Ben-Bassat and Dahan in 2012. Here the researchers investigate whether voters with a particular last name, which serves as a proxy for hamula (an Arab family based community) affiliation are more likely to vote for a candidate with the same last name as compared to other candidates. Strong empirical support is found for this hypothesis which means that candidates receive significantly more votes from voters with the same last name, which can be explained by SIT (Ben-Bassat & Dahan, 2012). In a research conducted by Lapinski and Mastro they found that even in jury decision making processes SIT seems to play a role. They found for example that

in rape cases men were more likely than women to believe the defendant was "not guilty" even in the strong evidence condition. Men also were more likely to assign shorter sentences to the defendant than were women which can be explained by SIT (Mastro & Lapinski). So there is quite some evidence for the influence of Social Identity theory in the decision making process of people. The FIFA Ballon d'Or award has also been used previously for research on SIT and decision making, Morsinkhof used the 2013 FIFA Ballon d'Or award to investigate whether voters did a preference for candidates that play on the position on the field (defender, midfielder or attacker) as them. Voters could see candidates who play on the same position as them a someone who is a part of their social group. Some evidence is found for this hypothesis but it's not strong enough to accept the hypothesis (Morsinkhof, 2014). This could be due to the fact that social groups are not formed naturally as opposed to the previously mentioned examples.

In addition to factors such as race and nationality, age is also a factor that can play a role in the group forming process of individuals. Tajfel and Turner used as age one the examples in their first formulation of SIT. Age group identification has been shown in the literature to have positive esteem consequences for its members. For example, young people may gain self-esteem through positively contrasting their youth with negative impressions of older adults (Harwood, Giles, & Ryan, 1995). Harwood also demonstrated that individuals display a preference for television shows that feature characters of their own age, even when all other aspects of content are controlled (Harwood, 1999)

## 2. Research method

### 2.1 Data

The data for the FIFA Ballon d'Or 2015 award is provided by the FIFA. The FIFA publishes the names of each voter, their role (captain of a team, coach, or journalist), country of origin as well as the ballot. Ages of voters were not provided so these have been collected through the internet, in most cases from Wikipedia pages of players. These were needed for the research question regarding analyses for groups based on age. For every player the voting deadline for the FIFA Ballon d'Or 2015 was the 20<sup>th</sup> of November 2015 (FIFA, 2016).

The voting system is as following, first members of FIFA's Football Committee and a group of experts from the publication *France Football* compile a shortlist of 23 players. After this journalists, national coaches and national captains from FIFA associated countries are given the chance to vote. The coaches, captains and media that vote for the award are asked to select their first, second and third choice for the award. Their first choice player gets 5 points, the second choice player is awarded three points while the third player receives 1 point. The player who receives the highest weighted percentage of all the votes is awarded the FIFA Ballon d'Or (Bailey, 2013). In 2015 the Argentinian player Lionel Messi was the winner of the award, Cristiano Ronaldo from Portugal finished in 2<sup>nd</sup> place while Brazilian forward Neymar Jr. finished in 3<sup>rd</sup> place.

### 2.2 Methodology

To test whether voters have a preference for candidates from their own region the voters are divided into two region based groups. The first group consist of voters from Europe, the second group consist of voters from Latin America. Voters from other regions are excluded for the analyses for this part, simply because these voters don't have any (or few) candidates from their own region. Of the 23 nominated players 22 are from either Europe or Latin America. In total 14 are from Europe and 8 hail from Latin America. The two historically dominant regions in football are Latin America and Europe, which has led to continuing debate on which of the two regions is really the 'king of football'. So to enhance the status of their respective groups we would expect that voters will prefer a player from their own region. Latin America is not an official

continent as opposed to Europe so there is some discussion on which countries are exactly part of it. For this thesis the countries that are apart of Latin America according to the Britannica Encyclopedia are used.

To test whether voters have a preference for players from their own region the sum of the voting points that every voter assigns to Latin American and Europe nominees is calculated. For example if a voter votes for Lionel Messi as his 1<sup>st</sup> choice (5 points), C. Ronaldo as a 2<sup>nd</sup> choice (3 points) and Neymar as his 3<sup>rd</sup> choice, the total number of points the he assigns to Latin American players would be 6 (Messi + Neymar) and 3 for Europe (C. Ronaldo). If the social identity theory holds we would expect that on average Latin American voters will assign more of their voting point to Latin American nominee than European voters and vice versa. To test whether this hypothesis holds a Mann Whitney test is carried out. This is a test can be used to compare two independent groups to test whether there is a differences between the two groups when the dependent variable is ordinal or continuous, but not normally distributed. Before determining which test to use a test for normality was used to determine whether the data was normally distributed or not (see Chapter 4: Results).

To examine the hypothesis that age plays a role in the decision making of voters only data of players is used. Ages of journalist were not available in the majority of the cases so these were left out of the analysis. Votes of coaches will also be left out, but this is for a different reason. In almost all cases the coaches are expected to be much older than the nominees they can vote on, which means that age identity thus cannot play a role in their voting choices. So the only votes that will be examined for the age identification part are those of the captains. To test this hypothesis the players are divided in three age based group. The three groups are captains under the age of 26, captains with a age between 26 and 30 and the last group consist of captains of 30 years old and above. These groups are based on historical data from all the FIFA World Cups which reveals that players tend to peak between the ages of 26 and 30. For example data reveals that of the 19 World Cup winning teams 16 of them had a average age of between 26 and 30 years (Carter, 2014). The exceptions were the Brazil team from 1963, which had a average age of 30.7. Two winning teams had a average of less then 26 these were the Uruguay team of 1950 (average age: 25.9) and the Argentinian team of 1978 (25.2), so

these were quite close to the interval. For each captain that was allowed to vote the mean age of the three candidates he voted on is calculated. If SIT holds we would expect that the mean age of the candidates a captain voted on will be higher when he is older. The statistical test that is used for this hypothesis is the Kruskal Wallis H. This is a rank-based nonparametric test that can be used to determine if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable (Laerd Statistics).



# Results

## 3.1 Test of normality

To determine which test to use a “test of normality” was carried out first to see whether the data used was normally distributed or not. The test used is a Shapiro-Wilk test, this test is suitable for sample of less than 2000 observations, as is the case here. The result can be seen in table 1. A significance level of 0,05 is used to decide whether or not to accept or reject the null hypothesis.

Table 1: Shapiro-Wilk ‘Test of normality’

		Statistic	df	Sig.
Point awarded to European players	Europe	,909	137	,000
	Latin America	,884	59	,000
Point awarded to Latin American players	Europe	,909	137	,000
	2,0	,884	59	,000

Both tests show a p-value of less 0.05, which means that the null hypothesis should be rejected. From this we can conclude that he data is not normally distributed, which means that we can’t use a parametric test. For not normally distributed samples the results

## 3.2 Europe vs. Latin America

The descriptive stats reveal that on average voters from EU and LA voter award 4,28 of their maximum of 9 voting point to European nominees, while for LA voters this number is considerably lower at 2,78. Which is quite a big difference and is in support of SIT. Because the of the fact that the test for normality test revealed that the data is not normally distributed a non-parametric test was used. The test used is a Mann Whitney test. The N for European voters is much larger than number of Latin American voters but this not an issue when using a Mann Whitney test.

**Table 2: Average voting points awarded to European candidates by regions**

	<b>N</b>	<b>Mean</b>
European voter's	137	4,28
Latin America voter's	59	2,78
Total	196	3,83

The Mann Whitney test results (table 2) show a significance level of 0.000 which is below the critical level of 0.05. Which means that the null hypothesis should be rejected and leads to the conclusion that there is indeed a difference in voting patterns between European and Latin American voters. From the mean from the descriptive stats we can determine that the difference is that on average European more of their voting points to European candidates than Latin American voters as was expected.

**Tabel 3: Mann Whitney test result**

	<b>Average voting points awarded to European players</b>
Mann-Whitney U	2469,000
Asymp. Sig. (2-tailed)	,000

If we run the same test again of we run the same test again but now for voting points received by Latin American players we see exactly the same results (table 4). The mean of voting points once again is much higher for voters from the same region as the candidate (6.22 vs. 4.72).

**Tabel 4: Average voting points awarded to Latin American candidates by region**

	<b>N</b>	<b>Mean</b>
European voter's	137	4,72
Latin American voters	59	6,22
Total	196	5,17

If we run the same test again of we run the same test again but now for voting points received by Latin American players we see exactly the same results (table 5).

**Tabel 5: Mann Whitney test results**

	Point awarded to Latin American players
Mann-Whitney U	2469,000
Asymp. Sig. (2-tailed)	,000

The p-value once again is 0.000, so the null hypothesis that there is no difference in rank between the groups is once again rejected.

From these results we can conclude that the average number of points that Latin American candidates received from voters from Latin America is higher than for European voters. Also both groups of voters on average award more voting points to Latin American candidates than to European candidates which is not surprising because of the fact that both 1<sup>st</sup> placed player and 3<sup>rd</sup> rank player from the results were Latin American. From the results of both the Mann Whitney test we can thus conclude that voters indeed are biased in their decision making toward players from their own region.

**Table 6: Top 10 from regional results (EU= Europe, LA= Latin America)**

Rank	Overall	EU voters	LA voters
1	Messi (LA)	Messi (LA)	Messi (LA)
2	C. Ronaldo (EU)	C. Ronaldo (EU)	C. Ronaldo (EU)
3	Neymar (LA)	Lewandowski (EU)	Neymar (LA)
4	Lewandowski (EU)	Neymar (LA)	Suárez (LA)
5	Suárez (LA)	Suárez (LA)	Müller (EU)
6	Müller (EU)	Müller (EU)	Mascherano (LA)
7	Neuer (EU)	Neuer (EU)	Neuer (EU)
8	Hazard (EU)	Hazard (EU)	J. Rodriguez (LA)
9	Iniesta (EU)	Iniesta (EU)	Lewandowski (EU)
10	Sánchez (LA)	Bale (EU)	Agüero (LA)

If we look at how the top 10 results would have looks for EU voters and LA voters in comparison to the overall results. This table supports the findings from the Mann Whitney U test, the result for European voters and Latin American voters are quite different than the overall results. In comparison to the overall results EU voters results s Europe only voting would have meant that a European player (Lewandowski) would have moved up to the 3<sup>rd</sup> place at the expense of a player from Latin America (Neymar). Furthermore a European player that did not end in the top 10 from the overall results (G. Bale) takes the 10<sup>th</sup> position at the cost of a Latin American player (A. Sanchez). For Latin American voters the differences are even bigger, 3 players from Latin America that didn't make the overall top 10 (Mascherano, J. Rodriguez and Aguero) would have ended in the top 10 at the expense of 2 European players and one Latin American player. One other noticeable difference is the position of Robert Lewandowski (Europe). Results from Latin American voters put him in the 9<sup>th</sup> position, while he ended on the 4<sup>th</sup> spot on the overall voting list. So in almost all of the cases candidates are better of when only voters from the same region as them would have voted. From the results of both the Mann Whitney test and the voting table differences we can thus conclude that voters indeed are biased in their decision making toward players from their own region. This bias seems stronger for voters from Latin America than for European voters.

### **3.3 Age based groups**

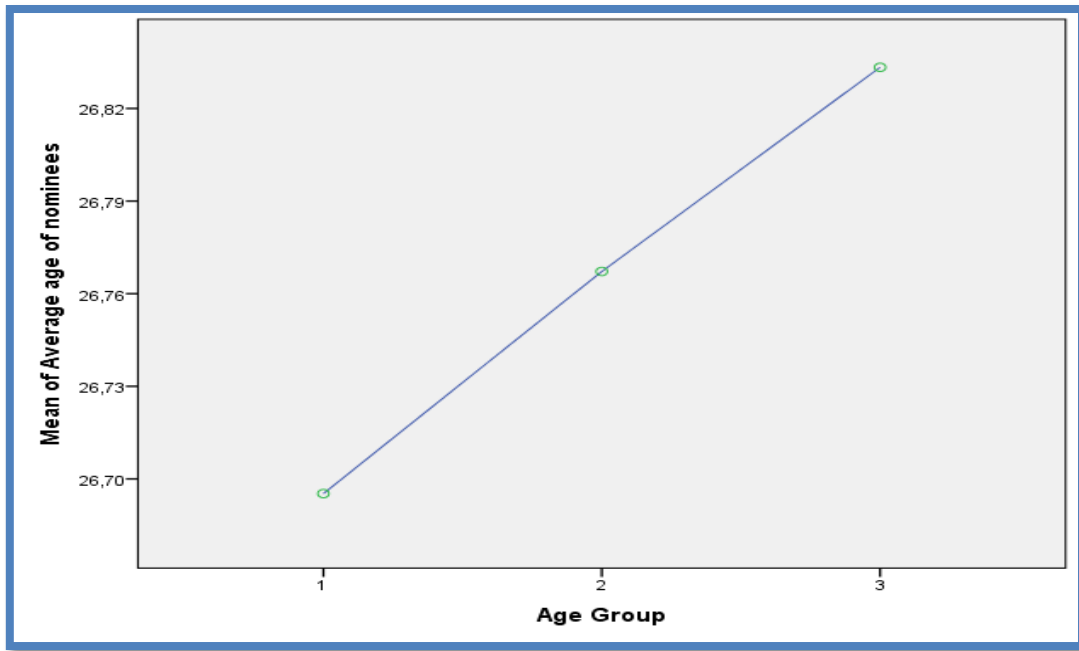
For this part the all the captains who had voting rights are classified into three age based groups:

Group 1: Captains younger than 26

Group 2: Captains between the age of 26 and 30

Group 3: Captains older than 30

When we look at the descriptive stats and the graph we can see that there indeed is a difference between the groups as was expected by SIT, although it must be noted that the difference is quite small. The difference between the average age of candidates of the youngest and oldest group is just 0.12 years. Older voters on average vote seem to vote for older candidates while for the youngest age group the mean age is the lowest.



A test of normality was carried out to determine which test could be used to test the hypothesis. The results (see Appendix A) showed that the data was not normally distributed which means parametric test can't be used. To test whether these mentioned findings were significant a Kruskal Wallis H test is used, which is a non-parametric test that can be used to see whether multiple means are equal or not. The Kruskal Wallis H test results don't back these findings, the p-value is much higher than 0.05 which means that the null hypothesis should not be rejected.

Table 6: Kruskal Wallis test results

	Average age of nominees
Chi-Square	,497
df	2
Asymp. Sig.	,780

In other words there is no difference in the means between the three groups, and there is no evidence that voters prefer a candidate from the same age or close to their own age.

## Conclusion

The main goal of this thesis was to determine whether the data would reveal any voting patterns that could be related to social identity theory, which would mean that voters not only base their decisions on performances on the pitch but that there are also some other factors which they consider when voting. The data was tested for region based and age and biases. The descriptive statistics for the region based groups indeed show that there are differences in voting patterns between voters from different regions. As expected by SIT voters show a strong bias for candidates from their own region. European voters prefer while Latin American voters prefer Latin American candidates, this effect seems to be stronger for Latin American voters than for European voters. For age based groups an Kruskal Wallis H test shows that there is no significant difference in voting patterns between the age based groups, although the average ages of the candidates are gradually higher for older groups. Overall the conclusion is that there are indeed other factors than performances on the pitch that voters consider when they voted for the FIFA Ballon d'Or 2015. On the one hand these results are expected when looking from a SIT, on the other hand the results could also be considered surprising because of the existence of a strong inner group rivalry between some of the Latin American countries. The results however do show that despite the inner group rivalry, the in-group candidates are still preferred over out-group (European) candidates.

### Discussion:

It must be noted that the difference between the voting patterns may have other reasons. For example it may be that voters vote on candidates from their own region because their countries are members of the same footballing federation (UEFA or CONMEBOL). This will have a positive influence on the chances of these countries for future continental based voting's. One example for such a voting is for example to determine which country gets to host the European Championships. One of the limitations of the data used is that this only allows analyses between Europe and Latin America, to make the case stronger analyses from other regions would have been useful to test whether regions that have no or less rivalry between them on a footballing level show region based biases or not. Unfortunately the data doesn't allow this because of the Latin American and European dominance in football. Home country bias may also

have played a role in the results, especially for the Latin American sample which has a relatively small number of observations. If for example 3 compatriots vote for a candidate as their first choice a candidate earn 15 points which is a considerable number especially for the lower half of the candidates. But the results don't only show differences in bottom half but also in the top 10 and even the top. So home country bias may have had a influence on the results, but the results can't be explained by home country bias only.

## Bibliography

- Bailey, R. (2013, December 9). *How Is the Ballon d'Or Winner Selected?* Retrieved August 4, 2016, from Bleacher Report: <http://bleacherreport.com/articles/1881792-how-is-the-ballon-dor-winner-selected>
- Ben-Bassat, A., & Dahan, M. (2012). Social Identity and Voter Turnout. *Public Choice*, 93-214.
- Carter, B. (2014, July 11). *When do footballers reach their peak?* Retrieved August 20, 2016, from BBC News: <http://www.bbc.com/news/magazine-28254123>
- Doise, W., Csepele, G., Dann, H., Couge, C., Larsen, K., & Ostell, A. (1972). An experimental investigation into the formation of intergroup representations. *European Journal of Social Psychology*, 202-204.
- Encyclopædia Britannica. (n.d.). *Countries of Europe*. Retrieved August 19, 2016, from Britannica: <https://www.britannica.com/topic/browse/Countries-of-the-World/Countries-of-Europe>
- Encyclopædia Britannica. (n.d.). *List of countries in Latin America*. Retrieved August 19, 2016, from Britannica: <https://www.britannica.com/topic/list-of-countries-in-Latin-America-2061416>
- FIFA. (2015). *Rules of Allocation*. Retrieved August 6, 2016, from FIFA: [http://resources.fifa.com/mm/document/ballon-dor/playeroftheyear-men/02/71/10/97/rulesofallocation2015-en\\_neutral.pdf](http://resources.fifa.com/mm/document/ballon-dor/playeroftheyear-men/02/71/10/97/rulesofallocation2015-en_neutral.pdf)
- FIFA. (2016, January 6). *All you need to know about the FIFA Ballon d'Or 2015*. Retrieved August 2016, 18, from FIFA: <http://www.fifa.com/ballon-dor/news/y=2015/m=12/news=all-you-need-to-know-about-the-fifa-ballon-d-or-2015-2741481.html>
- Harwood, J. (1999). Age Identification, Social Identity Gratifications, and Television Viewing. *Communication Reports*, 85-90.
- Laerd Statistics. (n.d.). *Kruskal-Wallis H Test using SPSS Statistics*. Retrieved August 11, 2016, from Laerd Statistics: <https://statistics.laerd.com/spss-tutorials/kruskal-wallis-h-test-using-spss-statistics.php>
- Mastro, D. E., & Lapinski, M. K. (n.d.). A Social Identity Approach to Understanding the Jury Decision Making Process: Race as a Social Indicator. *The Journal of Intergroup Relations*, 48-62.
- McKinley, C. J., Mastro, D., & Warber, K. M. (2014). Social Identity Theory as a Framework for Understanding. *International Journal of Communication*, 1049-1068.
- Morsinkhof, J. (2014). *The influence of social identity theory on voting*. Enschede.
- Sanderson, C. A. (2010). *Social Psychology*. New Jersey: John Wiley & Sons, Inc.
- Tafjel, H., & Turner, J. (1979). In W. Austin, & S. Worchel, *The Social Psychology of Intergroup Relations*. Monterey: Brooks-Cole.
- Yocco, V. (2014, July 25). *Social Influence: Incorporating Social Identity Theory Into Design*. Retrieved August 9, 2016, from Smashing Magazine: <https://www.smashingmagazine.com/2014/07/incorporating-social-identity-theory/>



## Appendix: A

### Shapiro-Wilk 'Test of normality': Age Identification

	Age Group	Statistic	df	Sig.
Average age of nominees	1	,887	35	,002
	2	,973	63	,189
	3	,960	64	,038