



**University of
Zurich^{UZH}**

Department of Business Administration

UZH Business Working Paper Series

Working Paper No. 391

Can Consumers See A Difference? An Experiment with High-Skilled Soccer Players

Carlos Gomez-Gonzalez, Helmut Dietl, David Berri, and Cornel
Nessler

First Version: 01 November 2021

Current Version: 15 November 2021

University of Zurich, Plattenstrasse 14, CH-8053 Zurich,
<http://www.business.uzh.ch/forschung/wps.html>



**University of
Zurich^{UZH}**

UZH Business Working Paper Series

Contact Details

Carlos Gomez Gonzalez

University of Zurich

Department of Business Administration

Plattenstrasse 14, CH-8032 Zurich, Switzerland

carlos.gomez@business.uzh.ch

Tel.: +41 44 634 53 14

CAN CONSUMERS SEE A DIFFERENCE?

AN EXPERIMENT WITH HIGH-SKILLED SOCCER PLAYERS ‡

ABSTRACT

Whether one looks at revenue, public and private investment, or media coverage; men's sports consistently do better than women's sports. Many people argue that these differences are driven by absolute differences in the quality of athletes in men's and women's sports. We begin by noting that absolute differences in athletic skill often do not drive demand in sports. We then move on to our primary research question: Can people truly see differences in men and women athletes? To answer this question, we use videos of professional women's and men's soccer. In some videos the gender of the athletes was clear to see. In other videos, though, the gender of the athletes was blurred. We find that participants only rate men's soccer videos higher when the gender of the players is visible. These findings reveal a bias in the evaluation of men's and women's soccer relevant for many other fields. Our results demonstrate that factors other than performance, e.g., social beliefs and stereotypes, have an important influence on how individuals evaluate women in sports.

‡ For their comments and suggestions, we would like to thank conference participants at the HSE Analytics for Management and Economics Conference (October, 2020), the 96th WEAI Annual Conference (June 27-July 1, 2021), the XI CIED Conference (June 2-4, 2021), the WEAI International Conference (March 17-19, 2021), and seminar participants at the NTNU Spring Seminar, Center for Sport and Culture Management Research (March, 2021).

1. INTRODUCTION

In December of 2019 it was announced that David Tepper had paid \$325 million for a new Major League Soccer (MLS) franchise in Charlotte, North Carolina (Negley, 2019). Just a few months later, it was thought the expansion fee for a new National Women's Soccer League (NWSL) team in Sacramento was between \$1 million and \$2 million (Kassouf, 2020). That same year, the Seattle Reign -- an existing NWSL team -- sold to French investors for a sale price that was less than \$4 million.

These disparities in private investment are also seen with respect to public investment. Matheson (2021) estimates that MLS has received \$1.8 billion in public subsidies from 1990 to 2020. These subsidies have been used to build various arenas for MLS franchises.

Prior to 2021, no one had ever built an arena for a NWSL team. That year, though, the owners of the NWSL team in Kansas City announced a new arena for their franchise that would cost \$70 million (Pruitt-Young, 2021). As those owners indicated, all of that funding was provided by the owners of the franchise. In other words, there was not public money dedicated to the building of the first arena dedicated to a NWSL team.

The disparities in private and public investment between the MLS and NWSL is difficult to explain when we consider the athletes the league employs. The NWSL employs many of the greatest players in women's soccer. In contrast, MLS players -- despite the "Major" name in the league's name -- is clearly a minor soccer league. Despite the hundreds of millions invested in the MLS - by both private and public entities -- no one would seriously rank this league among the best soccer leagues in the world.

Of course, even though MLS players can't compete consistently with players in the top European soccer leagues, there are those who think they can compete with NWSL players. U.S. Soccer -- which employs both MLS and NWSL players for the men and women's national teams -- stated in court documents in 2020 that "it is undisputed that the job of [Men's National Team] player requires materially more strength

and speed than the job of [Women's National Team] player." (Booker, 2020). Although U.S. Soccer later apologized for this statement, there is a sentiment that women are simply not as fast as men. And hence, this explains why women's soccer does not do as financially as well as men's soccer. In other words, because women cannot run as fast as men, women's soccer will never be as popular as men's soccer.

It certainly may be the case that women in soccer do run slower than men. But we do not think that is the question one should focus upon. For us, the more interesting question is whether or not those watching men and women play soccer can actually see a difference. In other words, our research seeks to understand the role difference in speed play in people's perceptions of women and men's soccer.

To address this question, we propose the following experimental design. Participants were shown videos of professional women's and men's soccer players scoring goals. Some of these videos were unmodified. In other videos, though, the gender of the players was blurred making it impossible for a viewer to know if they were watching men or women play soccer. Participants were then asked to evaluate the overall performance of the players. The results indicated that when participants could identify the gender of the players, male soccer players were rated higher. But when the gender of the participants could not be ascertained, participants thought women and men soccer players weren't different.

The approach we are adopting in the study of soccer players has been adopted in other research settings. For example, in the hiring process, research has shown that employers tend to discriminate less when the gender of the candidate is hidden (Åslund & Skans, 2012; Goldin & Rouse, 2000; Krause et al., 2012).¹ The concept of blinding identity features to improve decisions has also been explored in science, education, and criminal justice. For example, academic double blinded peer review enhances female authorship in journals (Budden et al., 2008) and the likelihood of project proposals led by female principal investigators to succeed in competitive calls (Johnson & Kirk, 2020). In education, blinded tests

¹ If no discrimination is present when hiring, Krause et al. (2012) find that blinding has no effect.

can be beneficial for pupils from minority groups (Hinnerich et al., 2015). Blinding procedures also inspired policy discussions to prevent racial biases of criminal prosecutors in the US (Sah et al., 2015).

To our knowledge, our study is the first to empirically analyze the effect of blinded procedures on gender biases in the context of professional sports.² The rest of the paper is structured as follows: Section Two discuss how sports fans often see absolute and relative differences in athletes. This is followed by a brief review of how women's sports is treated by the sports media. Section Four describes the method and data. Section Five then evaluates the results while Section Six concludes the paper with a discussion of the results.

2. ABSOLUTE VS. RELATIVE DIFFERENCES

Although most men would lose in a foot race against the top women athletes, it is quite clear -- as Table 1 indicates - that the very best male athletes are faster than the very best women. When it comes to the 100-meter dash, the best men are nearly a second faster than the best women.

[Table 1 near here]

² We should note, blinding has been used in the study of non-sports topics. We should also note that blinding identity features is not always effective to improve decisions or reduce biases. In hiring, gender or ethnicity-blind approaches to reviewing job applications do not directly improve the outcome of minority group members (Hiscox et al., 2017). Especially, if employers who are aware and willing to counteract biases are not able to directly observe the minority status of the applicant (Behaghel et al., 2015). The influence of blinded peer review in gender discrimination in grant and manuscript reviewing is often criticized (e.g., Ceci & Williams, 2011), and blinding the race of perpetrators when prosecutors examine their criminal cases has no statistically significant effect (Chohlas-Wood et al., 2021). The influence of blinding identity features in racial and gender outcomes in other settings (Chohlas-Wood et al., 2021) and joint evaluations (Bohnet et al., 2016) is yet to be fully understood.

It is important to note --as Table 2 indicates -- that men are not always faster than women. As the race gets longer, the differences between men and women get smaller and smaller. And at races beyond 195 miles, women actually run faster than men.

[Table 2 near here]

Despite what we see with respect to ultra-marathons, it does appear that men can sprint faster than women. But should that matter?

One might think absolute differences in athletes clearly drive differences in consumer demand in sports. But the sport of boxing suggests that sports fans seem more interested in relative differences.

Consider the "pound-for-pound" boxing rankings that various media outlets offer. For example, Table 3 presents ESPN's pound-for-pound rankings in July of 2021. Topping this list was Canelo Alvarez.

[Table 3 near here]

On November 6th of 2021, Alvarez fought Caleb Plant. As Coppinger (2021) noted when the fight was over: *"Canelo Alvarez entered the ring as the best pound-for-pound boxer in the world and the sport's biggest attraction. But now, for the first time in his career, he can call himself something else: undisputed champion."*

A month before the 5-9, 165 pound Alvarez entered the ring, Tyson Fury and Deontay Wilder fought for the WBC heavyweight title. Fury stands at 6-9 and weighs more than 270 pounds while Wilder is 6-7 and weighs more than 230 pounds. In sum, both Fury and Wilder literally tower over Alvarez. No one would seriously expect Alvarez to have a chance of defeating either man in a boxing match. In fact, the strict rules of boxing -- where fighters have to achieve the weight required by their boxing divisions before they fight -- prevent these fights from even taking place. Nevertheless, boxing fans are quite comfortable

arguing that Alvarez is a "better" fighter than either Fury or Wilder. In other words -- as Berri (2018) argued -- boxing fans rank relative skill over absolute skill in evaluating boxers.

Boxing is not the only place we see such a story. It is not uncommon in North America for athletic contests involving college teams to draw a larger audience than games involving professional teams. Most players on college teams will never play professional sports. Consequently, it is unlikely a college team would have much success against an entire team of professional players. Nevertheless, games between top college teams frequently draw a larger audience than many games involving professional teams.

3. THE VICIOUS CIRCLE IN WOMEN'S SPORTS

Of course, all of these examples come from the world of men's sports. When the subject of women's sports is raised, it seems suddenly the focus turns to absolute differences. This shift in focus illustrates that men and women are often judged by different criteria and that the criteria applied to women often employ negative gender stereotypes.

For example, it is argued that stereotypical judgments and misperceptions determine the slow professional advancement of women (Ellemers, 2018; Valian, 1999). Research also shows that negative gender stereotypes about women performance exist, e.g., in business (Heilman, 2001; Gupta & Turban, 2012), education (Boring, 2017; Mengel et al., 2019), politics (Dolan & Sanbonmatsu, 2009; Lawless, 2009), science (Moss-Racusin et al., 2012; Régner et al., 2019; Witteman et al., 2019), or sports (Burton, 2015; Darvin et al., 2018; Gomez-Gonzalez et al., 2019). In male-dominated occupations, women need to deal frequently with criticism of their hardiness and talent to succeed (Leslie et al., 2015; Morganson et al., 2010).

In the world of sports, the gender issues not only exist, but these difference very much seem to be magnified. As Travers (2008) notes, within sports hegemonic masculinity and gender differences are reproduced and naturalized. For example, the abilities of women are constantly and publicly questioned

by some fans and reporters on TV, Twitter, Facebook, and other social media outlets (Hyde, 2019; Scheadler & Wagstaff, 2018; Trolan, 2013). At least, that often happens when women are mentioned. As Cooney, et. al (2013) notes, often the media completely ignores women in sports. Whereas women are more than 40% of athletes, women only receive about 4% of the sports media's coverage.

A common assumption is that demand drives coverage and that the media provides what people prefer to consume (Gentzkow & Shapiro, 2006). There are certainly examples where men's sports events receive much greater television ratings than corresponding events in women's sports.³ But one can also argue that demand for men and women's sports isn't created in a vacuum. In fact, the coverage of women's sports often exhibits a vicious circle. Specifically, frequent exposure to stereotypical judgments about women may lead the sports media to provide less women's sports and many fans to demand less coverage. Hence, the way in which women are talked about in sports drives the supply and demand for women's sports.

Because of less media coverage, fans are not frequently exposed to women's sport and may infer lower quality and dislike. In addition to that, perceived physiological differences exacerbate the disadvantage for female athletes (Capranica et al., 2013; Chevront et al., 2005; Pallarés et al., 2012). Thus, because of inferred "lower quality", female sport is, supposedly, less interesting for viewers and less covered (Fink, 2015). One could make a similar argument for other sports, e.g., basketball, handball, hockey, or tennis.

³ For example, 1.12 billion people saw the men's world cup final and 260 million people saw the women's world cup final (see <https://digitalhub.fifa.com/m/2589b77c20849beb/original/njqsntrvdvqv8ho1dag5-pdf.pdf> and <https://digitalhub.fifa.com/m/5fd80f719fbff8e4/original/rvgxekduqpeo1ptbgcng-pdf.pdf>). We should note, though, it is not always the case men's sports do better than women's sports. According to U.S. Soccer (see <https://www.ussoccer.com/stories/2015/07/womens-world-cup-final-is-mostwatched-soccer-match-in-us-history>), the most watched soccer match in United States history was the 2015 women's World Cup.

4. METHOD AND DATA

One crucial part of the vicious circle we see in women's sports has never been empirically tested: Specifically, are performance differences between professional female and male athletes even noticeable for viewers? It is quite possible that men in soccer do indeed run faster than women. But it also entirely possible fans can't discern such differences.

4.1 Experimental videos

To test this idea, we randomly selected 10 videos from women's and men's club teams and national teams matches in 2019. The videos were between 5 and 14 seconds. We used goal plays that were chosen by either UEFA or FIFA as highlights for the season. We excluded videos that included long periods without movement. All videos showed the performance of high-skilled soccer players and contained some of the most watched goals on TV and social media. These videos showed only goal-scoring sequences. Table A1 in the Supplementary Information section gives an overview about the match, scoring player, and length of the videos.

To conceal the gender of the players, we manipulated every video frame. Supplementary Information 2 provides the technical information and the metrics of the manipulated videos in the software Adobe Premier Pro. We applied the same levels of mask feather, opacity, and expansion for each frame and video. We then selected the number of blocks for blurriness. We used the lowest level of blurriness that still allowed us to effectively conceal the gender of the players. Figure 1 shows an example of the two versions of the videos (original and blurred).

[Figure 1 near here]

We created two different groups; control and experimental. In the control group, participants evaluated a sequence of 10 videos, where nothing was blurred (similar to Figure 1 - L). In the experimental group,

participants evaluated the same sequence but with blurred videos (similar to Figure 1 - R). Thus, the only difference between control- and experimental group was the blurriness that conceals the gender of the players and stands. Every sequence consisted of five female- and five male soccer videos. We randomly allocated the position of each video to control for the influence that the position of the videos in the sequence have on the evaluation. Table A2 in the Supplementary Information section shows this allocation.

4.2 Participants and survey structure

We used Amazon Mechanical Turk (AMT), an online marketplace platform for workers, to recruit experiment participants. This platform is widely used because it makes it easier to create a more heterogenous sample, i.e., nationality and age, and the outcome does not significantly differ from traditional methods (Arechar et al., 2018; Snowberg & Yariv 2021). In the platform, workers exchange their labor for a monetary reward. Employers can specify additional qualifications that workers must meet to work on the task. Employers decide how much a worker will be paid for the task.

In our experiment, the task was to complete a survey - estimated to take 5 minutes. There were no qualification requirements for workers. The monetary reward was \$1.00. The workload was low, and the payment was high compared to common tasks on AMT.⁴

The payment was also sufficient to generate a significant sample of participants. Specifically, the sample consisted of 552 women and 674 men, who were on average 34 years old. 8 participants were from Africa, 20 from Asia, 450 from Europe, 644 from North America, 4 from Oceania, and 78 from South America.

⁴ Snowberg & Yariv (2021) used AMT for their research and, as a robustness check, halved the monetary incentive. Their results did not change.

Each of these participants were given a task labeled “Answer a survey about soccer”; with the description “Evaluate soccer performance (about 5 minutes)”. We added the keywords “survey; sport; soccer; video”. Before evaluating the videos, each participant was asked to provide personal information regarding age, gender, and country of residency. Additionally, participants were asked to answer the following questions with binary response possibilities: “Do you watch soccer?”, “Do you watch soccer daily (highlights or matches)?”, and “Do you prefer men’s or women’s soccer?”.

These questions were chosen because people who watch soccer regularly could show different behavior when evaluating women’s and men’s soccer. Additionally, research argues that behavior may be moderately influenced by explicit beliefs and values (Crandall & Eshleman, 2003; Devine et al., 2017). Therefore, we may find a relationship between the preference for women’s or men’s soccer and the evaluation of the groups, especially when the videos are unmodified (and the gender visible).

After every video we asked participants (1) to rate the performance of the players: “Overall, the performance of the players is” (1- Poor; 5- Excellent) and (2) to state the willingness to pay to watch the watch: “Would you pay to watch the match?”. The videos and questions were shown in subsequent survey screens and participants had to submit the final task in the last screen. Participants could take up to 20 minutes to complete the survey and go back and forth in the survey screens. We checked in a small subsample if asking questions before or after the videos influences the results. We did not find any differences.

5. RESULTS

Figure 2 shows the evaluation for individual videos. 9 out of 10 videos received a lower rating when they were blurred. Three unblurred men’s soccer videos received an average evaluation of above 4. No unblurred women’s soccer video received an evaluation of above 4. Most blurred videos received a rating around 3.5.

[Figure 2 near here]

In the next step, we combined men's soccer videos in one group (videos 1-5) and women's soccer videos in another group (videos 6-10). Participants rated men's soccer significantly higher when the videos were not blurred and the gender visible (4.012 vs. 3.839 average treatment effect (ATE) 0.173; Mann-Whitney $U, z = 3.429, P = 0.00, n = 580$). As shown in Figure 3, participants had no preference when the videos were blurred (3.583 vs. 3.556; ATE 0.027; Mann-Whitney $U, z = 0.846, P = 0.398, n = 646$).

[Figure 3 near here]

We tested if the results are robust when controlling for various covariates. Regression table 1 shows the responses for unblurred videos. We included all control variables stepwise. Model 1 presents only the evaluation of the videos. Model 2 includes participant demographics. Model 3 includes participant survey responses (note: the omitted response is that participants have no preference). Model 4 includes control variables for different sequences and participant country of residency. Most control variables have no statistically significant influence throughout the models. Only participants who watch soccer daily are more likely to rate all videos lower. Nonetheless, the control variables have no influence on the participants' preference for male soccer videos, which is consistent throughout all models.

[Table 4 near here]

Regression table 2 shows the responses for blurred videos. Under this experimental condition, participants no longer rate men's and women's soccer videos differently. This result is consistent across all models that include the controls. Participants that prefer men's soccer evaluated the blurred videos significantly lower. Additionally, females evaluated the blurred videos, in general, statistically significantly higher. We lose 84 observations in Model 3 as we did a coding error with the question regarding participants' preference for male or female soccer. Omitting these observations in Models 1 and 2 has no significant impact.

In Table A3 in the Supplementary information 2 section, we interact the variables *Female video* and *Gender of participant* to test if female and male participants evaluate unblurred videos differently. The results show that male participants evaluate female videos significantly lower than male videos. However, the difference between female and male participants evaluating female videos is not significant.

[Table 5 near here]

6. DISCUSSION

Many factors contribute to the disparity in the professional advancement of women (Ellemers, 2018; Valian, 1999). In many fields, women need to cope with stereotypes that questions their ability and the capacity to succeed (Leslie et al., 2015; Morganson et al., 2010), and their representation and visibility is severely limited. For example, in sports, the media does not cover men's and women's equally often (Cooky et al. 2013). In fact, women's sports receive only around 4% of all sports media coverage (UNESCO, 2021).

Some research has suggested that preferences drive the type of content displayed by the media (Gentzkow & Shapiro, 2006). Thus – it is argued -- fans prefer watching male sports because they derive a higher utility. There is reason to suspect, though, that fans are not evaluating men and women's sports by the same criteria. For example, fans of men's sports seem quite comfortable with evaluating male athletes relative to their competition. Absolute differences – as we see in boxing – are often not even considered. When it comes to comparing men and women, though, absolute differences are often cited.

Consequently, we hypothesize that gender stereotypes play an important role in defining preferences. Furthermore, we argue that the performance differences fans cite – such as the relative speed of men and women in soccer -- are not always visible. Our results demonstrate that differences between men and women performance are only significantly perceived when the gender of the players is visible. The

tendency to rate men's soccer higher completely disappears when the gender of the players is no longer distinguishable.

It is a common assumption that differences in physiology and professionalism between men's and women's soccer result in recognizable differences in quality. Our results show that this assumption must be reevaluated. This result is important for players, managers, and broadcasters alike. Men's soccer is omnipresent throughout societies, e.g., through commercials and social media. People are conditioned to see men and not women playing soccer. However, our results show that, at least, certain parts of women's soccer can effectively compete with men's soccer.⁵

Moderator factors regarding participant's demographics, preferences, and habits only have an influence on the overall evaluation of performance, but do not modify the gender bias. For example, participants that watch soccer daily rated both blurred and original videos lower than other participants. When people constantly watch soccer, they might be less impressed when they see professionals plays once more. Additionally, participants that prefer men's soccer rated blurred videos lower. Additionally, we do not find the expected influence of explicit beliefs and values on behavior (Crandall & Eshleman, 2003; Devine et al., 2017). Participants who explicitly express a preference for men's or women's soccer do not show a gender bias when evaluating the original videos. The quality criteria of fans who prefer women's and men's soccer may be closer than what many think, at least regarding the overall performance of players.

⁵ The soccer experience in the United States clearly illustrates that point. The U.S. women's soccer team has won multiple World Cups while the men often have trouble even qualifying to participate in the World Cup. In addition, in 2019 it was reported that U.S. women's soccer games were generating more revenue than games involving the U.S. men's soccer team (Johnson-Hess, 2019).

The findings from the experiment are robust and consistent but should be interpreted cautiously. This study has limitations that also provides opportunities to future research in this area. We focused exclusively on goal plays. However, a soccer match consists of various scenes that are interesting for viewers too (e.g., tackles, corners, or freekicks). Also, highlights are short. Many people enjoy watching whole matches. Finally, we cannot rule out the possibility that other variables explain better the observed relationship. For example, we did not control for the atmosphere in the stadium. Participants watched the videos without any sound as commentators frequently reveal the gender of the player. People often go to the stadium or watch a match not because of the quality but because of the atmosphere. This may also drive preferences for women's and men's soccer and differences in the evaluation of quality. These factors are missing from our analysis.

Additionally, the experimental design consists only of 10 videos because of the time needed to efficiently modify the videos and conceal the gender of the players (between two and six weeks each). A larger number of videos would ensure higher variability, which is desirable but not possible with respect to our workload. Moreover, we record the responses of participants at the individual level and outside the traditional context of sports consumption (computer-based). Our results cannot incorporate other factors such as social expectations, desired behavior, and group dynamics in other contexts.

In male-dominated fields, gender stereotypes are naturalized and reproduced. For example, the low ratings of media coverage for women's sports (Cooky et al. 2013; Travers, 2008), and the negative attitudes towards the performance of female athletes in some TV, radio, and social media (Hyde, 2019; Scheadler & Wagstaff, 2018; Trolan, 2013) contribute to the hegemony of men's sports and partly drive preferences of consumers. Our findings demonstrate that the evaluation of women's and men's athlete performance may be partly driven by gender prejudices and stereotypes.

The direction of this evidence should persuade consumers and sport's governing bodies to be mindful of biases when evaluating women's sports and making decisions about its promotion. Future research should

also help to shed light on the external factors beyond performance, e.g., stadium atmosphere, sounds, commentators, and group dynamics that may still make men's sport more attractive for a general audience.

REFERENCES

- Arechar, A. A., Gächter, S., & Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics*, 21(1), 99-131.
- Åslund, O., & Skans, O. N. (2012). Do anonymous job application procedures level the playing field? *ILR Review*, 65(1), 82-107.
- Behaghel, L., Crépon, B., & Le Barbanchon, T. (2015). Unintended effects of anonymous resumes. *American Economic Journal: Applied Economics*, 7(3), 1-27.
- Berri, D. (2018, June 13). Who is the best basketball player right now: Learning a lesson from boxing. *Forbes*. Retrieved from: <https://www.forbes.com/sites/davidberri/2018/06/13/who-is-the-best-basketball-player-right-now-learning-a-lesson-from-boxing/#8ce74760d52c>
- Bohnet, I., Van Geen, A., & Bazerman, M. (2016). When performance trumps gender bias: Joint vs. separate evaluation. *Management Science*, 62(5), 1225-1234.
- Booker, B. (2020, March 11). U.S. soccer apologizes for saying male players have "more responsibility" than women. *NPR. org*. Retrieved from: <https://www.npr.org/2020/03/11/814656567/male-players-have-more-responsibility-than-women-u-s-soccer-says-in-court-filing>
- Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, 145, 27-41.

- Brueck, H. (2020, January 7). Women are faster long-distance runners than men, and it's probably because they've got more estrogen. Business Insider. Retrieved from: <https://www.businessinsider.com/women-are-faster-long-distance-runners-estrogen-2020-1>
- Budden, A. E., Tregenza, T., Aarssen, L. W., Koricheva, J., Leimu, R., & Lortie, C. J. (2008). Double-blind review favors increased representation of female authors. *Trends in Ecology & Evolution*, 23(1), 4-6.
- Burton, L. J. (2015). Underrepresentation of women in sport leadership: A review of research. *Sport Management Review*, 18(2), 155–165.
- Capranica, L., Piacentini, M. F., Halson, S., Myburgh, K. H., Ogasawara, E., & Millard-Stafford, M. (2013). The gender gap in sport performance: equity influences equality. *International Journal of Sports Physiology and Performance*, 8(1), 99-103.
- Carter-Francique, A. R., & Richardson, F. M. (2016). Controlling media, controlling access: The role of sport media on Black women's sport participation. *Race, Gender & Class*, 23(1-2), 7-33.
- Ceci, S. J., & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences*, 108(8), 3157-3162.
- Chevront, S. N., Carter, R., DeRuisseau, K. C., & Moffatt, R. J. (2005). Running performance differences between men and women. *Sports Medicine*, 35(12), 1017-1024.
- Chohlas-Wood, A., Nudell, J., Yao, K., Lin, Z., Nyarko, J., & Goel, S. (2021). Blind justice: Algorithmically masking race in charging decisions. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 35-45).
- Cooky, C., Messner, M. A., & Hextrum, R. H. (2013). Women play sport, but not on TV: A longitudinal

study of televised news media. *Communication & Sport*, 1(3), 203-230.

Coppinger, M. (2021, November 6). Canelo Alvarez stops Caleb Plant to become undisputed middleweight champion. ESPN. Retrieved from: https://www.espn.com/boxing/story/_/id/32571503/canelo-alvarez-stops-caleb-plant-become-undisputed-super-middleweight-champ

Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin*, 129(3), 414-446.

Darvin, L., Pegoraro, A., & Berri, D. (2018). Are men better leaders? An investigation of head coaches' gender and individual players' performance in amateur and professional women's basketball. *Sex Roles*, 78(7), 455-466.

Devine, P. G., Forscher, P. S., Cox, W. T., Kaatz, A., Sheridan, J., & Carnes, M. (2017). A gender bias habit-breaking intervention led to increased hiring of female faculty in STEM departments. *Journal of Experimental Social Psychology*, 73, 211-215.

Dolan, K., & Sanbonmatsu, K. (2009). Gender stereotypes and attitudes toward gender balance in government. *American Politics Research*, 37(3), 409-428.

Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology*, 69, 275-298.

ESPN.com. (2021, June 23). Pound-for-pound rankings: A new No. 1, and a well-deserved debut in the top 10. ESPN. Retrieved from: https://www.espn.com/boxing/story/_/id/31509153/pound-pound-rankings-new-no-1-well-deserved-debut-top-10

Fink, J. S. (2015). Female athletes, women's sport, and the sport media commercial complex: Have we really "come a long way, baby"? *Sport Management Review*, 18(3), 331-342.

- Gentzkow, M., & Shapiro, J. M. (2006). Media bias and reputation. *Journal of Political Economy*, 114(2), 280-316.
- Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*, 90(4), 715-741.
- Gomez-Gonzalez, C., Dietl, H., & Nessler, C. (2019). Does performance justify the underrepresentation of women coaches? Evidence from professional women's soccer. *Sport Management Review*, 22(5), 640-651.
- Gupta, V. K., & Turban, D. B. (2012). Evaluation of new business ideas: do gender stereotypes play a role? *Journal of Managerial Issues*, 140-156.
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, 57(4), 657-674.
- Hinnerich, B. T., Höglin, E., & Johannesson, M. (2015). Discrimination against students with foreign backgrounds: Evidence from grading in Swedish public high schools. *Education Economics*, 23(6), 660-676.
- Hiscox, M. J., Oliver, T., Ridgway, M., Arcos-Holzinger, L., Warren, A., & Willis, A. (2017). Going blind to see more clearly: Unconscious bias in Australian Public Service shortlisting processes. Behavioural Economics Team of the Australian Government.
- Hyde, M. (2019, June 12). Pity the poor man who's had the Women's World Cup shoved down his throat. *The Guardian*. Retrieved from: <https://www.theguardian.com/football/blog/2019/jun/12/womens-world-cup-shoved-down-mens-throats>
- Johnson, S. K., & Kirk, J. F. (2020). Dual-anonymization yields promising results for reducing gender

- bias: A naturalistic field experiment of applications for Hubble Space Telescope time. *Publications of the Astronomical Society of the Pacific*, 132(034503), 1-6.
- Johnson-Hess, A. (2019, June 19). US women's soccer games now generate more revenue than men's—but the players still earn less. CNBC. Retrieved from: <https://www.cnn.com/2019/06/19/us-womens-soccer-games-now-generate-more-revenue-than-mens.html>
- Kassouf, J. (2020, August 19). Sacramento closes in on the NWSL, again. *The Equalizer*. Retrieved from: <https://equalizersoccer.com/2020/08/19/sacramento-republic-nwsl-expansion-2021-report/>
- Krause, A., Rinne, U., & Zimmermann, K. F. (2012). Anonymous job applications in Europe. *IZA Journal of European Labor Studies*, 1(1), 1-20.
- Lawless, J. L. (2009). Sexism and gender bias in election 2008: A more complex path for women in politics. *Politics & Gender*, 5(1), 70-80.
- Leslie, S. J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, 347(6219), 262-265.
- Matheson, V. (2021). Trends in public financing for professional sports stadiums in the US and Canada. Unpublished Working Paper.
- Mengel, F., Sauermann, J., & Zölitz, U. (2019). Gender bias in teaching evaluations. *Journal of the European Economic Association*, 17(2), 535-566.
- Morganson, V. J., Jones, M. P., & Major, D. A. (2010). Understanding women's underrepresentation in science, technology, engineering, and mathematics: The role of social coping. *The Career Development Quarterly*, 59(2), 169-179.
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science

faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, 109(41), 16474-16479.

Negley, C. (2019, December 16). Charlotte's MLS expansion fee reportedly a record \$325M -- two-thirds more than recent expansions. Yahoo Sports. Retrieved from: <https://sports.yahoo.com/charlotte-will-reportedly-become-30th-mls-franchise-for-record-325-m-two-thirds-more-than-recent-expansions-212710659.html>

Pallarés, J. G., López-Gullón, J. M., Torres-Bonete, M. D., & Izquierdo, M. (2012). Physical fitness factors to predict female Olympic wrestling performance and sex differences. *The Journal of Strength & Conditioning Research*, 26(3), 794-803.

Pruitt-Young, S. (2021, October 27). Kansas City team plans to build the first stadium created for women's league soccer. NPR. Retrieved from: <https://www.npr.org/2021/10/27/1049820541/kansas-city-team-plans-to-build-the-first-stadium-created-for-nwsl-womens-soccer#>

Régner, I., Thinus-Blanc, C., Netter, A., Schmader, T., & Huguet, P. (2019). Committees with implicit biases promote fewer women when they do not believe gender bias exists. *Nature Human Behaviour*, 3(11), 1171-1179.

Sah, S., Robertson, C. T., & Baughman, S. B. (2015). Blinding prosecutors to defendants' race: A policy proposal to reduce unconscious bias in the criminal justice system. *Behavioral Science & Policy*, 1(2), 69-76.

Scheidler, T., & Wagstaff, A. (2018). Exposure to women's sports: Changing attitudes toward female athletes. *The Sport Journal*, 19, 1-17.

Snowberg, E., & Yariv, L. (2021). Testing the waters: Behavior across participant pools. *American*

Economic Review, 111(2), 687-719.

Travers, A. (2008). The sport nexus and gender injustice. *Studies in Social Justice*, 2(1), 79-101.

Trolan, E. J. (2013). The impact of the media on gender inequality within sport. *Procedia-Social and Behavioral Sciences*, 91, 215-227.

UNESCO (2021). Gender Equality in Sports Media. Accessible at: <https://en.unesco.org/themes/gender-equality-sports-media>.

Valian, V. (1999). *Why so slow? The advancement of women*. Cambridge: The MIT Press.

Witteman, H. O., Hendricks, M., Straus, S., & Tannenbaum, C. (2019). Are gender gaps due to evaluations of the applicant or the science? A natural experiment at a national funding agency. *The Lancet*, 393(10171), 531-540.

Figure 1. Original Video (L) vs. Blurred Video (R)



Figure 2. Evaluation for All Videos. Two Experimental Conditions (Average Values)

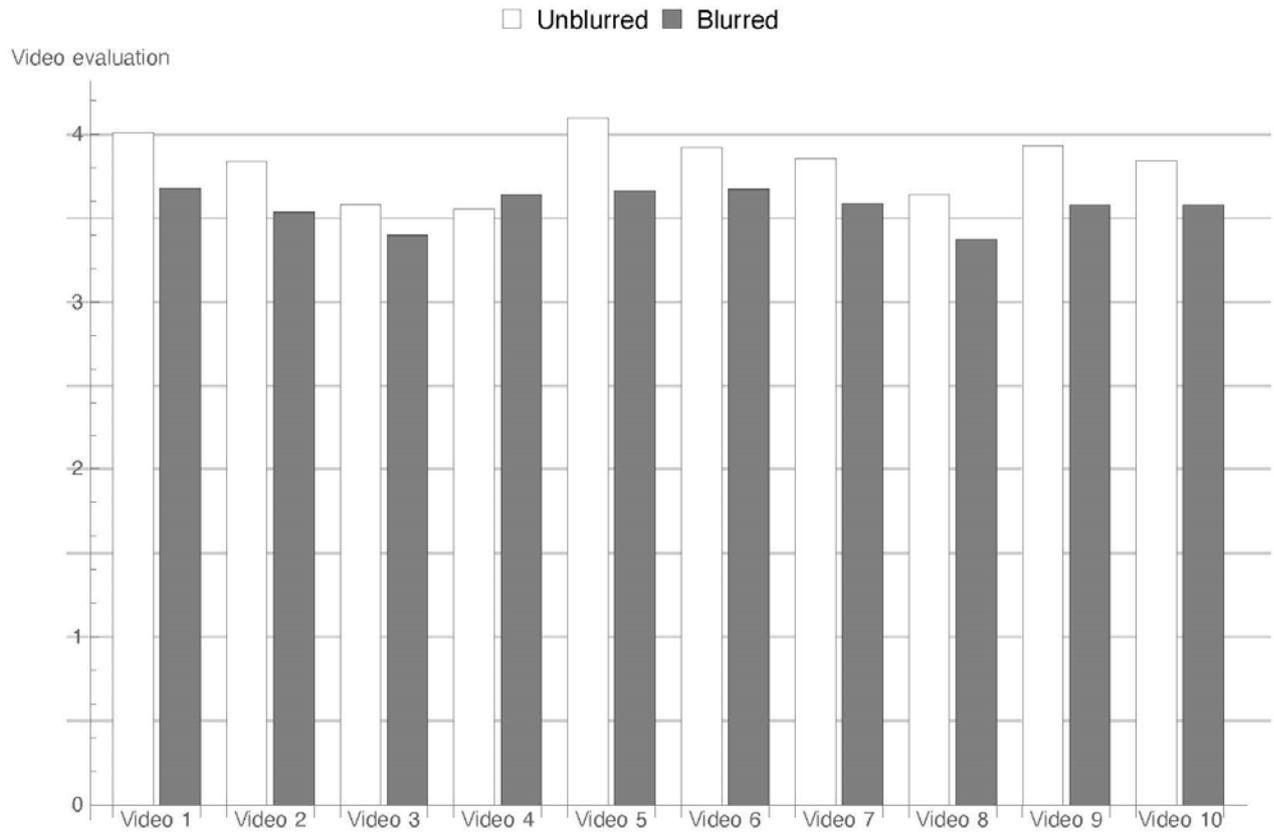


Figure 3. Differences in Evaluation Between Female and Male Soccer Videos. Two Experimental Conditions (Average Values).

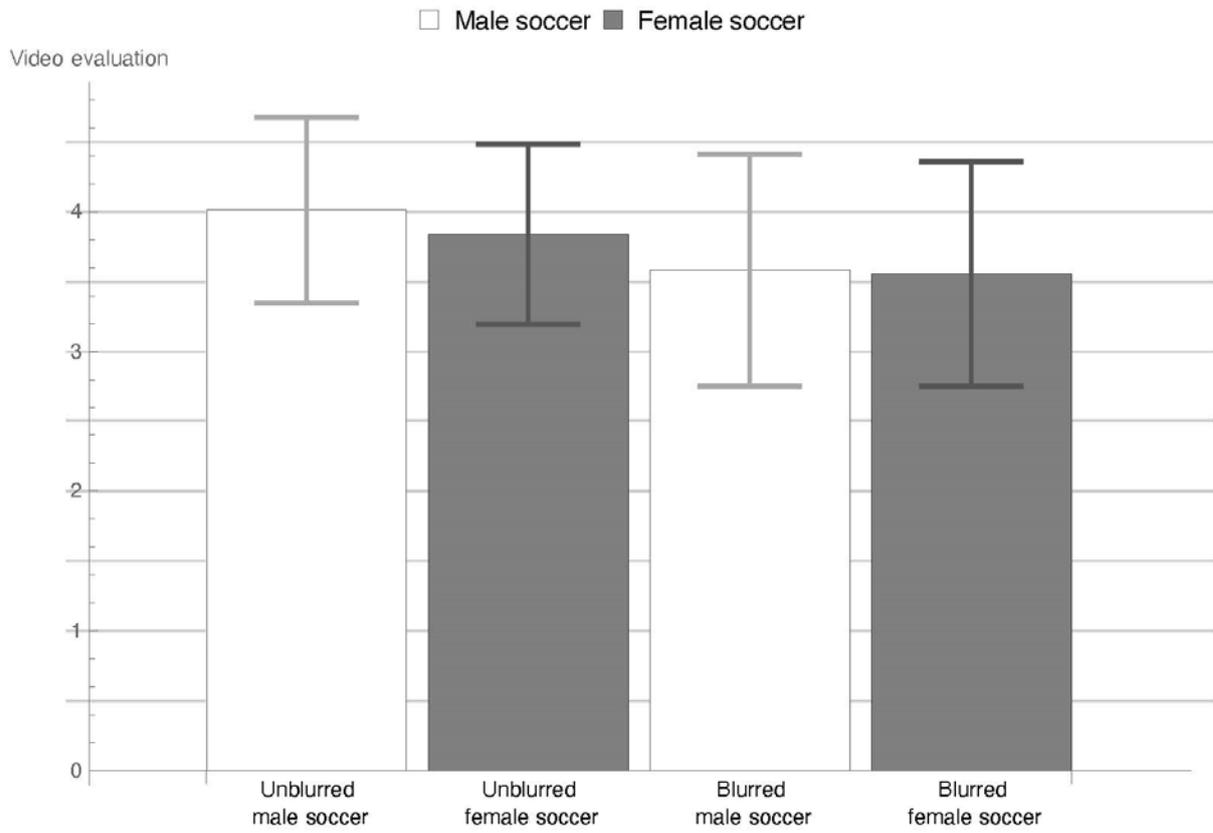


Table 1. Male vs. Female Sprinters

Rank	Male Sprinter	100-meter dash time	Female Sprinter	100-meter dash time
1	Usain Bolt	9.58	Florence Griffith-Joyner	10.49
2	Tyson Gay	9.69	Elaine Thompson-Herah	10.54
3	Yohan Blake	9.69	Shelly-Ann Fraser-Pryce	10.60
4	Asafa Powell	9.72	Carmelita Jeter	10.64
5	Justin Gatlin	9.74	Marion Jones	10.65

Note: Source a: <https://www.worldathletics.org/records/all-time-toplists/sprints/100-metres/outdoor/men/senior>. Source b: <https://www.worldathletics.org/records/all-time-toplists/sprints/100-metres/outdoor/women/senior>

Table 2. As the Race Gets Longer, Women Get Better!

Race	Percent men run faster than women
5K	17.9% faster
Marathon	11.1% faster
50K	5.3% faster
50 Miles	3.7% faster
100K	0.5% faster
100 miles	0.3% faster
195 miles	0.6% SLOWER

Note: Source: Brueck (2020)

Table 3. Pound-For-Pound Boxing Rankings from ESPN (July, 2021)

Boxer	Record	Boxing Division	Weight Limit in Division
Canelo Alvarez	56-1-2	Middleweight	165 pounds
Terence Crawford	37-0	Welterweight	147 pounds
Naoya Inoue	21-0	Bantamweight	118 pounds
Errol Spence, Jr.	27-0	Welterweight	147 pounds
Teofimo Lopez, Jr.	16-0	Lightweight	135 pounds

Note: Source: ESPN.com

Table 4: Regression Results. Responses to Unblurred Videos.

	Dependent variable: Video evaluation			
	Model 1	Model 2	Model 3	Model 4
Female video	-0.17*** (0.054)	-0.17*** (0.054)	-0.17*** (0.053)	-0.17*** (0.053)
Gender of participant (female)		0.038 (0.054)	0.035 (0.054)	0.080 (0.062)
Age of participant		0.031* (0.018)	0.028 (0.018)	0.029 (0.019)
Age ² of participant		-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Participant watches often			0.23* (0.14)	0.22 (0.14)
Participant watches daily			-0.30*** (0.073)	-0.26*** (0.073)
Participant has no preference			omitted	omitted
Participant prefers female soccer			0.079 (0.091)	0.13 (0.092)
Participant prefers male soccer			-0.059 (0.11)	0.015 (0.11)
Sequence control				Yes
Continent control				Yes
Constant	4.01*** (0.038)	3.40*** (0.34)	3.44*** (0.36)	3.36*** (0.60)
Observations	580	580	580	570
<i>Adj. R</i> ²	0.016	0.017	0.049	0.090

Note: a. The outcome variable is the evaluation of the overall performance of the players (1- Poor; 5- Excellent). b. Standard errors in parentheses. ***, **, * significance at 1, 5 and 10%, respectively; N = 580. c. Model 4 drops 10 participants who did not report place of residence.

Table 5: Regression Results. Responses to Blurred Videos.

	Dependent variable: Video evaluation			
	Model 1	Model 2	Model 3	Model 4
Female video	-0.027 (0.064)	-0.027 (0.063)	-0.023 (0.064)	-0.018 (0.059)
Gender of participant (female)		0.22*** (0.064)	0.19*** (0.065)	0.20*** (0.066)
Age of participant		0.033 (0.021)	0.031 (0.022)	0.024 (0.021)
Age ² of participant		-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Participant watches often			-0.042 (0.18)	-0.25 (0.17)
Participant watches daily			-0.26*** (0.086)	0.023 (0.085)
Participant has no preference			omitted	omitted
Participant prefers female soccer			-0.11 (0.10)	-0.081 (0.10)
Participant prefers male soccer			-0.57*** (0.13)	-0.35** (0.14)
Sequence control				Yes
Continent control				Yes
Constant	3.58*** (0.045)	2.64*** (0.40)	3.14*** (0.45)	3.13*** (0.52)
Observations	646	646	562	550
<i>Adj. R</i> ²	-0.001	0.036	0.090	0.186

Note: a. The outcome variable is the evaluation of the overall performance of the players (1- Poor; 5- Excellent). b. Standard errors in parentheses. ***, **, * significance at 1, 5 and 10%, respectively; N = 646. c. Model 3 drops 84 observations due to a coding error for the question about participants' preference for male or female soccer. Model 4 drops some additional participants who did not report place of residence. Omitting these observations does not significantly change the results in Models 1, 2 and 3.

Data availability

The data that support the findings of this study is publicly available in HarvardDataVerse, <https://doi.org/10.7910/DVN/P12G8I> (only accessible upon publication). We deleted individual identifiable information.

Author information

Affiliations

University of Zurich, Zurich, Switzerland

Carlos Gomez-Gonzalez & Helmut Dietl

Southern Utah University, Cedar City, United States

David Berri

Norwegian University of Science and Technology, Trondheim, Norway

Cornel Nesseler

Contributions

All authors contributed equally to this study and share the responsibility for its content.

Corresponding author

Correspondence to Carlos Gomez-Gonzalez.

Ethics declaration

The authors declare no competing interests.

Supplementary information**Table A1.** Videos in The Experiment.

Video	Match	Scoring Player	Length (in sec)	Female
1	Tottenham – Barcelona	Ivan Rakitić	14	No
2	Argentina - Croatia	Luka Modrić	6	No
3	Liverpool - Bayern	Sadio Mané	14	No
4	Manchester City - Donetsk	Raheem Sterling	9	No
5	Russia - Croatia	Denis Cheryshev	9	No
6	Chelsea - Lyon	Erin Cuthbert	6	Yes
7	Norway – England	Lucy Bronze	5	Yes
8	Netherlands – Sweden	Jackie Groenen	7	Yes
9	Nigeria – Korea Republic	Asisat Oshoala	13	Yes
10	United States - Thailand	Alex Morgan	5	Yes

Table A2. Allocation of Female and Male Videos in Subsequent Screens (S) Across Sequences (Seq.)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Seq. 1	♂	♂	♀	♂	♂	♀	♀	♂	♀	♀
Seq. 2	♀	♀	♀	♀	♂	♂	♂	♂	♂	♀
Seq. 3	♀	♂	♀	♂	♀	♂	♀	♂	♀	♂
Seq. 4	♀	♂	♀	♂	♂	♂	♀	♂	♀	♀
Seq. 5	♀	♂	♀	♀	♂	♂	♂	♀	♀	♂
Seq. 6	♀	♀	♀	♀	♂	♂	♂	♂	♂	♀
Seq. 7	♂	♀	♀	♀	♂	♀	♂	♂	♀	♂

Notes: ♀ female video; ♂ male video

Table A3. Regression Results. Responses to Unblurred Videos. Interaction Terms

	Dependent variable: Video evaluation			
	Model 1	Model 2	Model 3	Model 4
Male video # Male participant	omitted	omitted	omitted	omitted
Male video # Female participant	0.0059 (0.077)	-0.021 (0.077)	-0.0071 (0.076)	0.044 (0.082)
Female video # Male participant	-0.21*** (0.076)	-0.21*** (0.076)	-0.21*** (0.075)	-0.21*** (0.074)
Female video # Female participant	-0.12 (0.077)	-0.15* (0.077)	-0.14* (0.076)	-0.092 (0.082)
Participant has no preference		omitted	omitted	omitted
Participant prefers female soccer		0.026 (0.090)	0.079 (0.091)	0.13 (0.092)
Participant prefers male soccer		-0.15 (0.11)	-0.059 (0.11)	0.015 (0.11)
Age of participant		0.026 (0.018)	0.028 (0.018)	0.029 (0.019)
Age squared of participant		-0.00031 (0.00024)	-0.00034 (0.00023)	-0.00035 (0.00023)
Participant watches often (1/0)			0.23* (0.14)	0.22 (0.14)
Participant watches daily (1/0)			-0.30*** (0.073)	-0.26*** (0.073)
Sequence control				Yes
Continent control				Yes
Constant	4.01*** (0.054)	3.54*** (0.34)	3.46*** (0.36)	3.38*** (0.60)
Observations	580	580	580	570
R^2	0.020	0.034	0.063	0.121

Note: Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Supplementary Information 2. Parameters of The Manipulated Videos in Adobe Premiere Pro.

Mosaic function parameters:

- Feature details:
 - o Stands and Fans (Surroundings):
 - Mask feather: 10.0; Mask Opacity: 100.0 %; Mask Expansion: 0.0 – 10.0
 - Horizontal Blocks: 10; Vertical Blocks: 10
 - Sharp Colors: No
 - o Players:
 - Mask feather: 10.0; Mask Opacity: 100.0 %; Mask Expansion: 0.0 – 10.0
 - Horizontal Blocks: 100; Vertical Blocks: 100
 - Sharp Colors: Yes

Lumetri Color function parameters:

- Feature details:
 - o “Black” Team:
 - HSL Secondary
 - Key
 - “Set Color” and “Add Color” with pipette in Video (select surroundings of players); Invert Mask: Yes
 - Correction:
 -



- Temperature: -100.0
- Tint: 100.0
- Contrast: 100.0
- Sharpen: 0.0
- Saturation: 0.0

- o “White” Team:
 - HSL Secondary
 - Key
 - “Set Color” and “Add Color” with pipette in Video (select surroundings of Players); Invert Mask: Yes
 - Correction:
 -



- Temperature: 100.0
- Tint: -100.0
- Contrast: 100.0
- Sharpen: 0.0
- Saturation: 0.0