Article

PERCEPTION

Finding Neymar: The Role of Colour in the Detection and Discrimination of Football Kits

Perception
2021, Vol. 50(7) 615–626

DOI: 10.1177/03010066211019370
journals.sagepub.com/home/pec
SAGE

Liam Burnell
University of York, UK; Mindflick, UK; Liverpool John Moores University, UK

Peter Thompson
University of York, UK

Abstract
Association Football (hereafter football) is a fast-moving sport in which rapid decisions need to be made; where are other players, are they on my team? Two experiments investigated how kit variations affect the search of teammates. Experiment 1 confirmed that discriminability is slower when playing in crossed kits (e.g., red shirts-blue shorts vs. blue shirts-red shorts), versus uniform kits (e.g., all red vs. all blue). Experiment 2 found that there is significant confusion when both teams wear the same-coloured shorts. Based on these findings, we suggest changes to the Laws of the game concerning kit colours.

Keywords
visual search, football, visual detection, perception, kit colour, shirts

Date received 7 March 2020; accepted: 16 November 2020

Introduction

When a footballer looks up to make a pass to a teammate, he is faced with a visual search task. First, he must be able to detect players against the background. Second, he must be able to discriminate players of his own team from those of the opposition. Law 4 of the Game states (amongst other things) that "The two teams must wear colours that distinguish

Corresponding author:
Liam Burnell, Mindflick, 2 The Barn, Hathersage, Hope Valley, S32 1BB, UK.
Email: liam@mindflick.co.uk

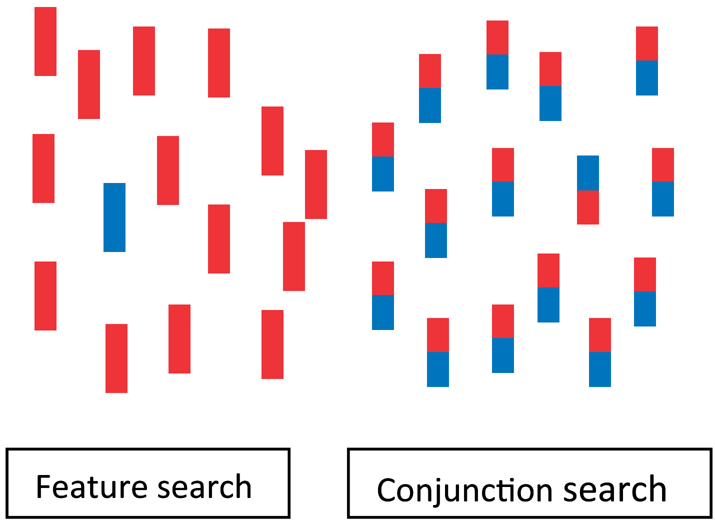

Figure 1. A comparison of feature search (left), versus conjunction search (right). See text for details. *Note.* Please refer to the online version of the article to view the figure in colour.

them from each other and the match officials." However, there is no further guidance as to what constitutes this "discrimination."

The discrimination of teammates from opposing players is a visual search task, which requires the identification of a target amongst a field of distractors (see Figure 1). Whilst a complete review of this literature is unnecessary here, it is worth considering the principles outlined by Feature Integration Theory (Treisman & Gelade, 1980).

This theory outlines two forms of search: Feature and Conjunction. Feature search is parallel, pre-attentive and "pop-out" in nature, with search focusing on one particular feature (e.g., colour or orientation). Alternatively, conjunction searches require focal attention to combine multiple feature maps together to identify a target. Under this form, search time increases linearly as a function of set size, meaning search is more effortful. Therefore, avoiding discrimination tasks of this nature in football may be advantageous for players and officials, alike.

In an attempt to show the relevance of such principles to football, Georgeson et al. (2005) compared the discrimination of configurations made up of single colours (uniform conditions, e.g., all blue vs. all red) against those of an overlapping nature (crossed conditions; e.g., red shirts, blue shorts vs. blue shirts, red shorts - see Figure 2). Rather than having to indicate the position of one target amongst many distractors, as in a standard visual search task, participants had to indicate how many of the target configurations were in each scene as quickly and as accurately as possible. Performance was largely accurate across conditions, but response times increased dramatically when kits were crossed, suggesting a more inefficient form of search. However, the schematic nature of the stimuli makes us uncertain of the relevance of the findings to the task facing players in a real game.

To see whether these findings extend to the task facing players on the football pitch, experiment 1 attempted to extend such findings by using stimuli more akin to the confusion that footballers (and officials) may face. One such instance may be a Southampton goal that was disallowed incorrectly when the assistant referee had to make an offside call in conditions akin to a conjunction search; a situation which may have made the decision harder than it needed to be—see Figure 3.

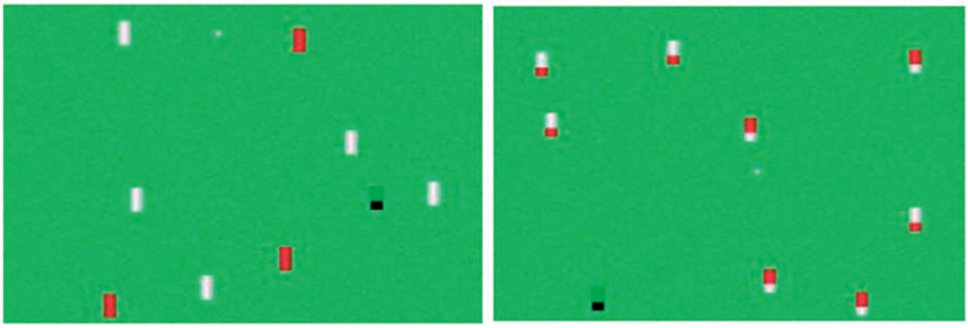

Figure 2. Samples of (left) the uniform, and (right) the crossed stimuli used by Georgeson et al. (2005).
Note. Please refer to the online version of the article to view the figure in colour.

Figure 3. Controversy in the 2017 EFL Cup Final, played with crossed kits; a context which is likely to have made discrimination of teams harder (retrieved from Hackett, 2017; © Action Images).
Note. Please refer to the online version of the article to view the figure in colour.

Experiment 1

Georgeson et al. (2005) showed discrimination of targets from distractors is slower when stimuli are "crossed" (e.g., red shirt-blue shorts vs. blue shirt-red shorts), relative to when uniform (e.g., all red vs. all blue). Experiment 1 attempted to replicate this finding in a more realistic setting.

To measure discrimination, participants had to identify the number of players wearing a target kit within the scene. The proportion of correct responses and average response time measured performance as a function of condition (uniform or crossed), target shirt colour (red or blue), and number of targets (3, 4, or 5). A high proportion of correct responses was expected throughout, whilst response times were hypothesised to increase significantly under crossed conditions.

Methods

An opportunity sample of 16 participants completed the study. All reported normal or corrected-to-normal vision and had no history of any visual problems, including colour blindness. A within-subjects design assessed the effect of condition (uniform kits or crossed kits), target shirt colour (red or blue), and the number of target players (3, 4 or 5) on performance. Proportion of correct responses and response time assessed discriminability.

Figure 4. An example of (A, left) uniform, and (B, right) crossed stimuli used in Experiment 1
Note. Please refer to the online version of the article to view the figure in colour.

A Dell E173FPf, 17″ monitor presented stimuli using PsychoPy v.1.84.2 (Pierce, 2007). To create the uniform stimuli (shown in Figure 4A), cropped images of players wearing all red (Liverpool FC) and all blue (Leicester City FC) were taken from in-game footage of the PlayStation4 game FIFA17. They were placed on an achromatic image taken from the centre of the pitch at White Hart Lane, on May 8, 2016 (Mills, 2016). This experiment focuses on the effect of kit colours, and so to eliminate any possible detection problems caused by various colours within the crowd or the pitch we have opted for a completely achromatic background. There were always eight outfield players in each image, all scaled for size based on the size of the average player (1.8 m) compared with the height of the goalposts in the image (1.1 cm in the image; usually 3.05 m tall, approximately 50 m away from the camera). Four players were positioned on the edge of the centre circle (3.52 cm tall; 9.1 m from the camera), and four players scaled to appear 24 to 27 m away from the viewer (1.19 cm–1.33 cm). Each scene also contained a goalkeeper (0.71 cm; 45 m away) and a referee (1.78 cm; 18 m away).

Crossed stimuli (red shirts, blue shorts vs. blue shirts, red shorts; shown in Figure 4B) were created by altering the shorts and sock colour of the same players used for the uniform stimuli. To do this, socks and shorts of the blue and red kits were coloured using a red (RGB = 161, 12, 26) and blue hue (RGB = 21, 45, 164), respectively. This hue was similar to that of the uniform kits. The same background image was used, with players scaled as per the uniform stimuli. The display, viewed at 57 cm, subtended 25° horizontally by 13.5° vertically.

Figure 5 depicts an individual trial; a fixation cross was followed by one of the images described in the materials. Participants were able to move their eyes about the scene and had to indicate whether three, four, or five target players were in the image, using the appropriate number key on a keyboard, as quickly and as accurately as possible. As soon as they had made their response, the next trial began.

Participants completed a short practice block, before four blocks of experimental trials took place. Before each block, the target shirt colour appeared on the screen. This varied between blocks, along with condition (uniform or crossed), with order counterbalanced between participants. The number of target players varied randomly trial-to-trial, with each possible number of targets displayed 8 times, resulting in 24 trials per block.

After completing all four blocks, debriefing took place.

Results and Discussion

The response times, calculated based on all completed trials—see Figure 6A—show speed of response was almost twice as long under crossed conditions versus uniform, regardless of

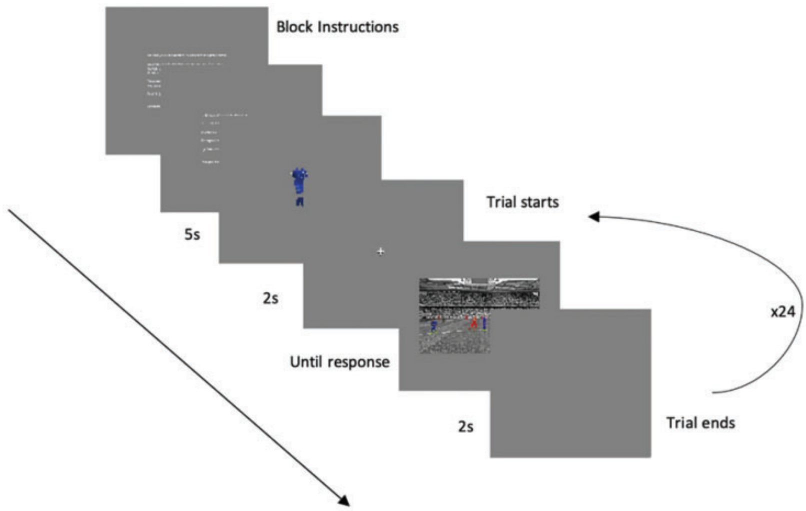

Figure 5. Illustration of the procedure of Experiment 1, with block instructions followed by a series of trials. *Note*. Please refer to the online version of the article to view the figure in colour.

target shirt colour. Figure 6B also shows response time increased as the number of targets increased, with a seemingly linear increase under crossed conditions.

A repeated measures analysis of variance (ANOVA) assessed the effect of condition (uniform, crossed), target shirt colour (red, blue), and number of targets (3, 4, and 5) on response times. This found response times to significantly increase under crossed conditions, versus uniform; $F(1, 15) = 234.54$, $p < .001$, partial $\eta^2 = 0.94$, and as the number of targets increased, $F(2, 30) = 19.98$, $p < .001$, partial $\eta^2 = 0.57$. However, no significant effect of target shirt colour was found, $F(1, 15) = 0.22$, p = .643, partial $\eta^2 = 0.02$.

A significant interaction between number of targets and condition, $F(2, 30) = 8.36$, $p = .001$, partial $\eta^2 = 0.36$, further quantified these effects, with the increase in response time as number of targets increased being driven by crossed conditions (compared with uniform). Specifically, repeated contrasts revealed this increase in response time under crossed conditions to be approaching significance as the number of targets increased from 3 to 4 ($p = .06$), and highly significant as the number of target kits increased from 4 to 5 ($p = .004$). This suggests that whilst response times for uniform search remained fairly stable, response times for crossed kits increased in a fairly linear fashion—consistent with conjunction search.

No significant effect was found in the interactions between Condition × Target Shirt Colour, $F(1, 15) = 0.00$, $p = .951$, Target Shirt Colour × Number of Target, $F(2, 30) = 1.70$, $p = .157$, and Condition × Target Shirt Colour × Number of Targets, $F(2, 30) = 1.60$, $p = .219$.

With regard to accuracy, the accuracy of responses was high in all conditions. When conditions were uniform, accuracy was always over 90%. However, when conditions were uniform, accuracy dropped slightly to 87% when the number of targets was 5—see Figure 7.

A repeated measures ANOVA revealed accuracy to be significantly lower for crossed kits versus uniform kits, $F(1, 15) = 30.34$, $p < .001$, partial $\eta^2 = 0.67$, and accuracy to decrease as

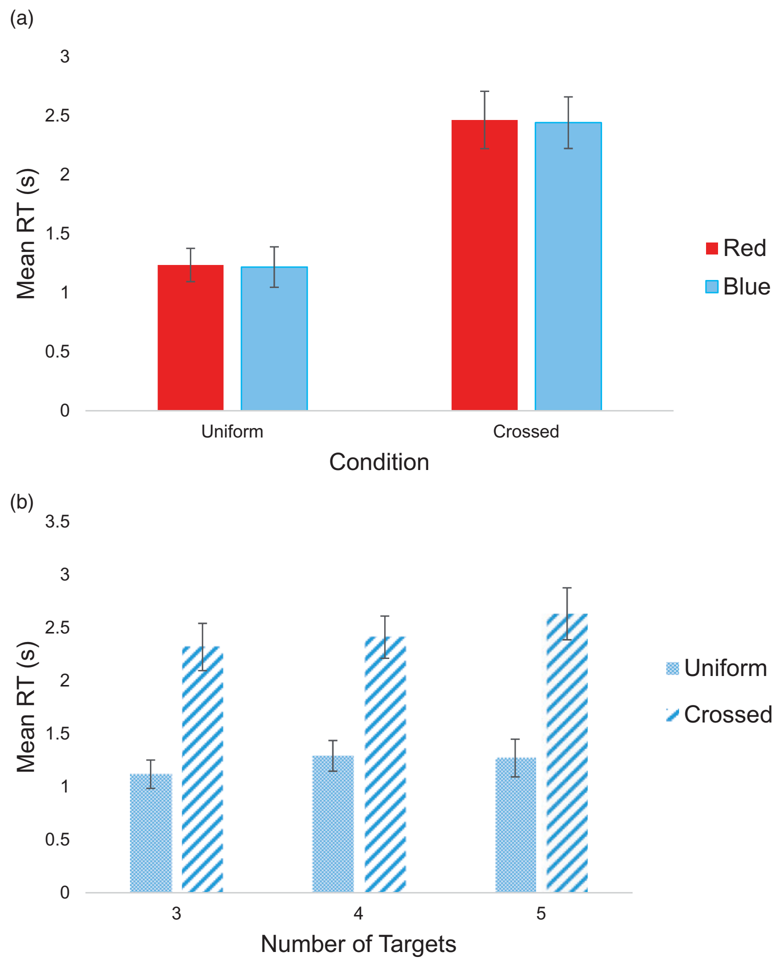

Figure 6. Mean RT to detect target players (with 95% confidence intervals), as a function of (A) condition and target shirt colour, and (B) number of targets and condition. RT: response time.

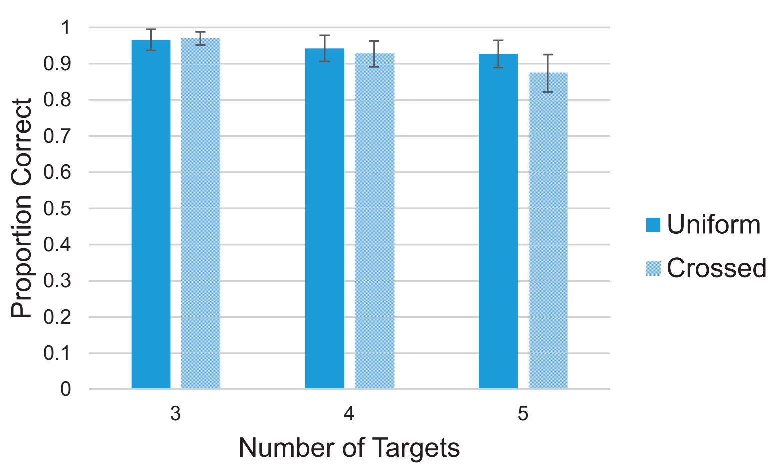

Figure 7. A bar chart showing proportion of correct responses (with 95% confidence intervals), as a function of number of targets and condition.

the number of targets increased, $F(2, 30) = 7.21$, $p = .003$, partial $\eta^2 = 0.31$. However, there was no effect of target shirt colour, $F(1, 15) = 3.62$, $p = .077$, partial $\eta^2 = 0.19$.

A significant interaction between condition and number of targets quantified such effects, $F(2, 30) = 13.32$, $p < .001$, partial $\eta^2 = 0.47$, with Helmert contrasts revealing this interaction to be driven by the larger decrease in accuracy under crossed conditions (vs. uniform conditions) as the number of targets increased from 4 to 5, $F(1, 15) = 18.44$, $p = .001$.

There were no significant interaction effects between Condition × Target Shirt Colour, $F(1, 15) = 2.67$, $p = .123$, Target Shirt Colour × Number of Targets, $F(2, 30) = 2.16$, $p = .133$, and Condition × Target Shirt Colour × Number of Targets, $F(2, 30) = 0.66$, $p = 523$.

These results illustrate that whilst responses were largely accurate, crossed conditions may impair the discrimination of two teams, especially as the number of target players increases. This is particularly noticeable when looking at response time, with discrimination taking around twice as long under crossed conditions. This replicates the findings of Georgeson et al. (2005), and the seemingly linear increase in response time under crossed conditions seems to be in line with search of a conjunction nature.

This shows that search may be seriously compromised when teams wear crossed kits, as player identity depends on a conjunction search. This perhaps explains the problems faced by the assistant referee in the 2017 English Football League (EFL) Cup Final (Figure 3).

Experiment 2

Law 4 states that "the two teams must wear colours that distinguish them from each other and the match officials." Experiment 1 has shown us that conjunction search slows down decision making, indicating that different shirt colours alone are not sufficient, and that shorts colour can also be important, as is seen when teams wear crossed kits.

There is no requirement for teams to wear different shorts colours as can be seen in Figure 8. In this example, the team in red shirts/white shorts may be able to identify a teammate with a feature search for "red." Meanwhile, the team in white shirts/white shorts may require a conjunction search of "white + shirt" to identify a teammate.

Figure 8. Image showing kits worn in the game between Swansea City and Middlesbrough in the English Premier league played on April 2, 2017. Note both teams are playing in white shorts (© Premier League, 2017; retrieved from https://www.premierleague.com/match/14338)

Note. Please refer to the online version of the article to view the figure in colour.

Experiment 2, using the same methods as Experiment 1, investigated whether wearing a common shorts colour would disadvantage players in such a way during the visual search task that they face.

In particular, we investigated two matches:

Match 1

R B
R v R

Team 1 in all red plays Team 2 in blue shirts and red shorts

Here, we predict that response times will be long for Team 1 as they must carry out a conjunction search of red + shirt. For Team 2, we predict a faster response as the target is a feature search for blue. The shorts colour being red should not hinder the search. Therefore, in this match, Team 1 is at a disadvantage.

Match 2

R B
R v G

Team 1, the all red team, now faces Team 3 with blue shirts and green shorts.

The task for Team 1 is now a simple feature search for red. Therefore, we predict they will have faster response times than in Match 1. This comparison of the response times of Team 1 in the two matches will show whether shorts colour does matter.

In this second match, Team 3 still faces an all-red opposition, as did Team 2 in the first match, but now their green shorts avoid a clash in shorts colour. This search is still a feature search for blue (or green, if the player decides to search for the colour of the shorts), and so Team 3's task here is similar to that of Team 2, though possibly a little easier by virtue of having a lower degree of target-distractor similarity than Team 2 did in Match 1.

It is possible that in this match, Team 1 will have an advantage over Team 3 as they have a larger surface area of red, than Team 3 has surface area of blue (or green).

Participants

An opportunity sample of 14 participants completed the study. All reported normal or corrected-to-normal vision and had no history of any visual problems, including colour blindness.

A within-subjects design assessed the effect of condition (Match 1: shorts of the same colour as the opposition; or Match 2: shorts of a different colour from the opposition), target shirt colour (red or blue) and the number of target players (3, 4 or 5) on performance. Proportion of correct responses, and response times assessed discriminability.

Results and Discussion

Again, response times across all trials were used as an indicator of how effortful search was, and indicated a detrimental effect on response times for teams wearing the same colour shorts (Match 1), compared with when shorts colour was different (Match 2)—Figure 9.

To confirm such an effect, a repeated measure ANOVA assessed the impact of kit similarity (same shorts, different shorts), target shirt colour (red, blue), and number of targets (3, 4, and 5) on response times. It showed response times to be significantly longer when shorts were similar, versus different; $F\,(1, 13) = 65.72$, $p < .001$, partial $\eta^2 = 0.84$, and as the

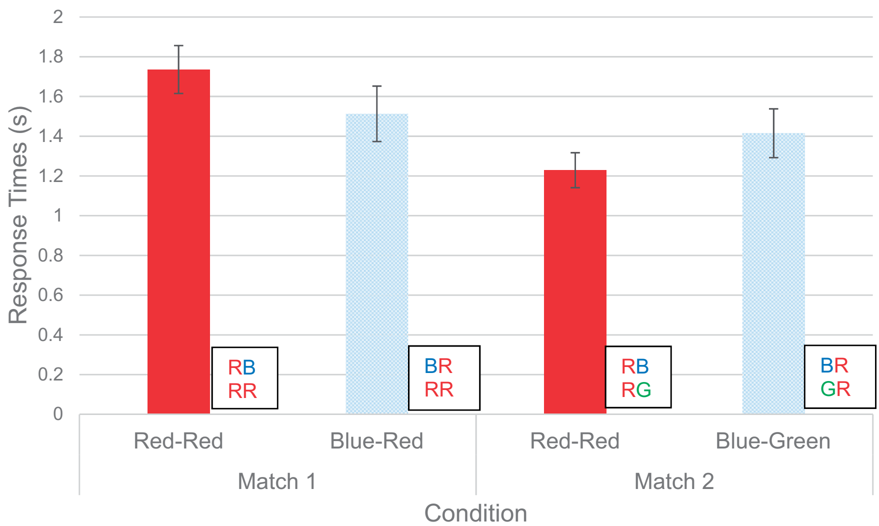

Figure 9. Bar chart showing the time to identify target kits in the presence of distractor kit in Match 1 (Column 1: time to identify red/red vs. blue/red; Column 2: blue/red vs. red/red) and Match 2 (Column 3: red/red vs. blue/green; Column 4: blue/green vs. red/red).

number of targets increased, $F(2, 26) = 13.09$, $p < .001$, partial $\eta^2 = 0.50$. However, there was no significant main effect of shirt colour, $F(1, 13) = 0.13$, $p = .729$.

Elaborating on the above effects, a significant interaction effect between kit similarity and colour, $F(1, 13) = 61.03$, $p < .001$, partial $\eta^2 = 0.82$, illustrated that whilst the removal of an overlap in shorts colour reduced response time for both teams (red shirts—$p < .001$; blue shirts—$p = .012$), this decrease in response time was larger for the team wearing red shirts. So much so, when an overlap in shorts colour was present (i.e., Match 1), the team in red shirts and red shorts took longer than the team wearing blue shirts and red shorts to respond—thus being disadvantaged ($p = .009$). However, when the overlap in shorts colour was removed (i.e., Match 2), the advantage was reversed, with the team in red shirts and red shorts now at an advantage compared with the team wearing blue shirts and green shorts ($p < .001$). There were no other significant interaction effects.

Furthermore, accuracy of response was also high (90% or more) across all conditions, and a repeated measures ANOVA assessed the effect of kit similarity (same shorts, different shorts), target shirt colour (red, blue), and number of targets (3, 4, and 5) on accuracy.

There was no significant effect of kit similarity, $F(1, 13) = 1.26$, $p = .281$, partial $\eta^2 = 0.09$, and shirt colour, $F(1, 13) = 0.33$, $p = .577$, partial $\eta^2 = 0.03$. However, there was a significant effect of number of targets, $F(1, 13) = 3.50$, $p = .045$, partial $\eta^2 = 0.21$, with the accuracy of response decreasing slightly as number of targets increased. Furthermore, no significant interaction effects were found between any of the reported variables, with the interaction effects of Kit Similarity $\times$ Target Shirt Colour, $F(1, 13) = 1.80$, $p = .203$, Kit Similarity $\times$ Number of Targets, $F(2, 26) = 0.05$, $p = .953$, Target Shirt Colour $\times$ Number of Targets, $F(2, 26) = 2.02$, $p = .153$, and the three-way interaction, $F(2, 26) = 1.50$, $p = .243$, all failing to reach significance.

These results regarding accuracy seem to indicate that as the number of targets increases, accuracy decreases across all four conditions. Taken in combination with response time data, this suggests that as the number of targets increased within this experiment, search became increasingly hard across all conditions.

Together, these results partially confirm the predictions above—the all-red team is disadvantaged by the opposing team's red shorts in Match 1, leaving the team in blue shirts and red shorts at a slight advantage. This is likely to be because the all-red team's task is a conjunction search (red + shirt), compared with the easier feature search of "just blue" for the opposing team. Thus, the red-shirted Middlesbrough may have gained an advantage over the all-white Swansea, Figure 8, because of the unique colour within their kit (find the red).

Meanwhile, in Match 2, where no overlap in shorts colour is present, those wearing all one colour now seem to be at a slight advantage, compared with those wearing shorts which are a different colour to their primary shirt colour. This may be because of the larger surface area of red present in the kit, prompting a faster search, although further investigation may be needed to confirm such claims.

Finally, it is worth noting that quicker response times occur for both teams when there is no overlap in shorts colour. This is unsurprising for the all red team, as their task has gone from a conjunction to a feature search. However, for the blue shirted team, their search in both matches is just a feature search for "blue." Therefore, it may be that this slight improvement experienced by the blue shirted team is driven by the decrease in target-distractor similarity experienced in Match 2, due to the removal of the overlapping colour (Pashler, 1987). Alternatively it could just reflect that in Match 2 either blue or green denotes a target whereas in match 1 only the blue shirt is a target.

Altogether, these results suggest that shorts colour does affect search on the football pitch—and should be considered when making decisions about clashes in kit—particularly when there is potential for opposing teams to wear the same colour shorts.

General Discussion

These experiments attempted to unpick how kit colour affects the discriminability of teammates in association football. Experiments 1 and 2 expanded on research conducted by Georgeson et al. (2005), confirming that the discrimination of players in crossed kits, which demands a conjunction search, is much more effortful than when players are wearing uniform kits. Response times double under such conditions and, as shown in Experiment 2, are also affected by instances of overlapping shorts colour.

Despite investigation of only red and blue kits in this article, Georgeson et al.'s original study, which investigated red-white, blue-white, and red-blue crossed colours, suggest such impairments occur regardless of the colours involved. This could have implications for the laws of the game, with current guidelines only prompting a change in kit when the referee believes there to be a direct overlap in shirt colour. With crossed conditions also impairing the ability to discriminate two teams quickly and accurately, rules that prevent this from occurring may be beneficial in preventing misplaced passes, and incorrect refereeing decisions. Making discrimination easier with such rule changes, which should also provide guidance for the avoidance of overlapping shorts colours (Experiment 2), may also help players to avoid making misplaced passes, such as that made by Michael Bostwick, who played an inch-perfect pass to a steward after mistaking him for a teammate in a 2016 league game (BT Sport, 2016).

Whilst we have investigated the discriminability of one kit against another, there remains the problem of how discriminable a team kit may be against the general background of the ground. We have preliminary data that suggest that a grey kit is hard to see. This supports comments made by Sir Alex Ferguson, following his side's defeat to Southampton in 1996. United had won 11 of their previous 12 games and Southampton were threatened with relegation. Playing in an all-grey away kit, United were three-nil down at half time. Ferguson's first words to the team were reportedly, "Get that kit off, you're getting changed" (Sharpe, 2006). In the second half, United, wearing blue and white, got one goal back but still lost the game 3–1. The kit change was punished by the Football Association with a fine of £10,000, described by Ferguson as "the best £10,000 I ever spent" (Duxbury, 1996).

Similar reports have supported the case for grey being a particularly poor colour in promoting detectability, with the then Wales manager, Chris Coleman, stating before a World Cup qualifier against Austria in 2016, "I can tell you no-one wants to wear our grey kit. I will be happy if I never see that kit again" (Pritchard, 2016). These reports suggest that a grey strip does not "pop-out" from the background, a problem which may be exacerbated when the team strip is close in colour to the stadium seating and a crowd decked out in team colours—particularly relevant when games take place in front of empty stadiums. Comments by Derek Adams in 2017, the then Plymouth Argyle manager, reflect this point (BBC Sport, 2017), as he stated,

> We have changed the kit about a bit—had white shorts and white socks because of the reason that the kit is too dark . . . It blends in with the seating and it blends in with the grass, and that's the reason we have changed it for a number of home games this season.

Further investigation may be warranted in light of such comments.

To expand the applicability of these findings, we realise it would be helpful to translate them on to in-game decisions. Therefore, it may be beneficial to develop paradigms that simulate in-game circumstances, such as the one used by Olde Rikkert et al. (2015), in which participants had to make decisions about moving players and their position against a marker.

Conclusions

These experiments begin to explore the complex world of football kits, an area of research that has received an underwhelming amount of coverage. Results reveal that discrimination is harder when players wear crossed kits, rather than uniform kits, due to the conjunction-like search they demand. This discrimination task also seems to be made harder when opposing teams wear the same colour shorts. Therefore, our results suggest that teams might reconsider their kit designs and that changes to the Laws of the game might prevent some sources of confusion in player identification.

Acknowledgements

The authors thank Jonny Davies, Tottenham Hotspur's Media relations manager, for supplying the photograph of the White Hart Lane pitch used in these experiments. The picture is now available at https://tommills.co.uk/gigapixel/tottenham-hotspur-gigapixel/. The authors would also like to extend our thanks to our anonymous reviewers for their helpful comments to improve the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/ or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article

ORCID iD

Liam Burnell https://orcid.org/0000-0002-3284-1352

References

BBC Sport. (2017). Derek Adams: Plymouth wore away kit 'for fans', not because of clash with grass. *BBC Sport*. https://www.bbc.co.uk/sport/football/39204434

BT Sport. (2016). *Watch as kit confusion sees Peterborough player pass to steward*. https://sport.bt.com/football/watch-as-kit-confusion-sees-peterborough-player-pass-to-steward-S11364114716110

Duxbury, N. (1996). United drop grey strip after black day. *Independent*. http://www.independent.co.uk/news/united-drop-grey-strip-after-black-day-1305058.html

Georgeson, M., Lampard, J., & Georgeson, J. (2005). Kits, colours, and confusion: A pilot study of vision and football. *Perception*, *34*, 633–637. https://doi.org/10.1068/p5302

Hackett, K. (2017). League Cup final offside error will haunt Stuart Burt—and possibly Southampton—for eternity. *Telegraph*. http://www.telegraph.co.uk/football/2017/02/26/league-cup-final-offside-error-will-haunt-stuart-burt-possibly/

Mills, T. (2016). Tottenham hotspur gigapixel panorama. *Tom Mills Photography*. https://tommills.co.uk/gigapixel/tottenham-hotspur-gigapixel/

Olde Rikkert, J., Haes, V. D., Barsingerhorn, A. D., Theelen, T., & Olde Rikkert, M. G. (2015). The colour of a football outfit affects visibility and team success. *Journal of Sports Sciences*, *33*, 2166–2172. https://doi.org/10.1080/02640414.2015.1064156

Pashler, H. (1987). Target-distractor discriminability in visual search. *Perception & Psychophysics*, *41*, 285–292. https://doi.org/10.3758/BF03208228

Pierce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pritchard, D. (2016). Chris Coleman on Wales' away kit: 'I'd be happy if I never saw it again'. *BBC Sport*. http://www.bbc.co.uk/sport/football/37513212

Sharpe, L. (2006). 13.04.96 Manchester United's grey day at The Dell. *The Guardian*. https://www.theguardian.com/football/2006/apre/15/sport.comment2

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136. https://doi.org/10.1016/0010-0285(80)90005-5