International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Role of Colour Priming on Memory and Sustained Attention: A Direct Replication Study

Dona Matthew[1], Koteshwar Ramesh Rakesh[2]

[1]Graduate, Department of Psychology, CHRIST (Deemed to be University)
[2]Assistant Professor, Department of Psychology, CHRIST (Deemed to be University)

Abstract
Colour priming is hypothesized to enhance memory and sustained attention through emotional arousal. This study conducted a direct replication of Dzulkifli and Mustafar (2013) in an Indian university population. Sixty undergraduate students (ages 18–25) were randomly assigned to three conditions: warm colour (red; n = 20), cool colour (blue; n = 20), or white control (n = 20). Participants completed a digit memory span task (MSPAN) and a sustained attention task (TOAV). Contrary to the original hypothesis, Kruskal-Wallis tests revealed no significant effects of colour priming on memory span, $\chi^2(2) = 0.808$, $p = .668$, $\eta^2 = .02$, or sustained attention reaction time, $\chi^2(2) = 4.998$, $p = .082$, $\eta^2 = .053$. Effect sizes were negligible. These null findings align with recent meta-analytic evidence of publication bias in colour psychology and suggest that colour priming effects may not generalize universally across populations. The replication contributes to cumulative psychological knowledge by testing boundary conditions and reducing publication bias.

Keywords: colour priming, memory, attention, replication, null results, publication bias

1. Introduction

Colour is a ubiquitous feature of human experience and has long been suspected to influence cognition. The theoretical proposition that colours, particularly warm colours such as red enhance attention and memory performance has generated substantial empirical research over the past four decades. However, recent meta-analytic evidence raises critical questions about the robustness and replicability of these effects [1]. The present study conducts a direct replication of Dzulkifli and Mustafar's (2013) influential investigation of colour priming on memory and sustained attention, extending this work by examining a novel population (Indian university students) and including methodological boundary condition extensions. This replication attempts to contribute to the cumulative knowledge base on colour–cognition relationships and to assess the generalizability of prior findings across cultural contexts.

1.1. Colour Perception and Cognition: The Theoretical Framework

Colour is fundamentally a perceptual and cognitive phenomenon. Unlike the electromagnetic radiation that constitutes the physical basis of colour, human colour perception involves complex neural processing, information from the retina is transmitted through millions of retinal neurons to the brain's visual cortex, where it is interpreted and integrated with prior knowledge and emotional associations. Beyond its perceptual properties, colour has been theorized to influence cognitive processes through two primary mechanisms. First, colour can modulate emotional arousal and affective state, with warm colours (red, orange, yellow) associated with heightened activation and energy, while cool colours (blue, green) are

associated with calm and cognitive focus [2, 3]. Second, colour can function as an attentional cue, capturing and directing visual attention to colour-salient regions in the visual field, which in turn may enhance encoding of those regions into memory [4]. These dual mechanisms namely, affective arousal and attentional capture, form the theoretical foundation for the hypothesis that colour priming enhances cognitive performance.

1.2 Empirical Foundations

Early research established compelling findings linking colour to memory performance. Dzulkifli and Mustafar (2013), in a widely-cited study, demonstrated that coloured backgrounds significantly enhanced recognition memory for geometric shapes compared to monochrome backgrounds, and that this effect was larger for warm colour conditions than cool colour conditions. Their work suggested that colour priming i.e., background colour presentation preceding or accompanying a cognitive task, could enhance memory retention through emotional arousal [4]. Similarly, Olurinola and Tayo (2015) showed that colour choices in educational materials can increase encoding and retrieval success, with particular benefits for warm-colour stimuli [5]. Greene, Bell, and Boyer (1983) established that warm colours (red, orange, yellow) elicit stronger attentional engagement than cool colours (blue, green), with downstream consequences for memory consolidation [6]. Collectively, these foundational studies argued for a functional role of colour in enhancing attention and memory, leading to widespread application of colour principles in educational and commercial design.

1.3 The Replicability Crisis and Publication Bias in Colour Psychology

Despite these early positive findings, recent evidence has revealed a troubling pattern in colour psychology research: publication bias and variable replicability. A comprehensive 2020 meta-analysis examining the effect of red colour on cognitive performance across 73 experiments found that while published studies reported a mean effect size of $d = -0.40$, this effect approached zero after applying corrections for publication bias [1]. This suggests that studies reporting null or weak effects are less likely to be published, inflating the apparent size of colour effects in the literature. Additionally, several landmark replication attempts have failed to reproduce earlier findings, including studies examining colour priming effects on attention [7] and colour-induced emotional effects on task performance [8]. Cave and colleagues (1996) reported that task-relevant stimulus properties (shape, brightness, edge contrast) override colour priming effects in many cognitive tasks, suggesting that colour effects may be context-specific and task-dependent rather than universally applicable [9]. These replicability challenges highlight a critical gap: the generalizability and robustness of colour priming effects remain uncertain, particularly across different populations, task contexts, and measurement approaches. This uncertainty is especially pronounced for non-Western populations, where colour psychology research is limited.

1.4 The Need for Replication in Indian Populations

Colour psychology research has been predominantly conducted in Western, educated, industrialized, rich, democratic (WEIRD) populations. Cultural factors including symbolic associations with colours, educational practices, and exposure histories, may moderate colour effects on cognition [10]. To date, no published studies have systematically examined colour priming effects on memory and sustained attention in Indian university populations, creating both a theoretical and practical gap. Indian educational institutions represent a diverse population with distinct cultural backgrounds, learning practices, and cognitive profiles. Examining whether colour priming effects replicate in this context is therefore both scientifically important (testing boundary conditions for colour effects) and practically valuable (informing educational design in Indian universities).

1.5 The Present Study: Direct Replication with Methodological Extensions

The present study was designed as a direct replication of Dzulkifli and Mustafar (2013), with methodological extensions to examine boundary conditions. Our specific aims were to: (1) test whether warm colour (red) priming enhances memory span compared to cool colour (blue) and neutral (white) control conditions; (2) test whether colour priming enhances sustained attention performance; and (3) examine the generalizability of these effects to an Indian university population using a convenience sample.

We employed several methodological enhancements relative to our implementation context: we used a validated computerized task battery (Psychology Experiment Building Language; PEBL), which has demonstrated reliability for measuring memory and attention in young adult populations [11, 12]. We employed a three-group design comparing warm colour (red), cool colour (blue), and white control conditions, allowing us to distinguish between colour-specific effects and general effects of background colour presentation. Given recent evidence of publication bias and small effect sizes in colour psychology, we also planned analyses of effect sizes and conducted post-hoc power estimation to contextualize our findings.

1.6 Hypothesis and Predictions

Based on the theoretical framework outlined above and the original findings of Dzulkifli and Mustafar (2013), we hypothesized that warm colour (red) priming would enhance both memory span and sustained attention performance compared to cool colour (blue) and white control conditions. Specifically, we predicted: (1) participants in the red (warm colour) condition would demonstrate higher digit span scores (MSPAN) compared to blue and white conditions; (2) participants in the red condition would show faster reaction times and fewer errors on sustained attention tasks (TOAV) compared to blue and white conditions; and (3) the magnitude of colour effects, if present, would be quantifiable via effect size estimation to allow comparison with prior literature.

However, in light of recent replication failures and meta-analytic evidence of publication bias, we remained alert to the possibility of null or minimal effects, which would constitute a scientifically important finding contributing to the literature on replicability in cognitive psychology [10].

2. Method

2.1 Research Design

This study employed a between-groups quasi-experimental design with three conditions: warm colour priming (red), cool colour priming (blue), and control (white background). The design was modelled after Dzulkifli and Mustafar (2013), with the addition of an explicit white control condition as a boundary condition extension. The dependent variables were memory span performance and sustained attention performance, measured via validated computerized tasks from the Psychology Experiment Building Language (PEBL) battery. This study was designed as a direct replication attempt; as such, we explicitly aimed to test the generalizability of Dzulkifli and Mustafar's (2013) findings to an Indian university population using convenience sampling.

2.2 Participants and Sampling Procedure

2.2.1 Sample Composition

The final sample comprised 60 undergraduate students (age range: 18–25 years, $M = 21.3$ years, $SD = 1.8$) from CHRIST (Deemed to be University), Delhi NCR, India. The sample included students from multiple

disciplines and represented both male (n=35; 58.34%) and female (n=25; 41.66%) participants.

2.2.2 Sampling Method and Recruitment

We employed convenience sampling, recruiting participants through course announcements in the Department of Psychology, departmental bulletin boards, and email communications to registered undergraduate students. While convenience sampling limits statistical generalizability to the broader population, this approach is standard in replication research and allows for explicit definition of the population sampled (i.e., CHRIST University undergraduates). Participants self-selected into the study based on availability and interest.

2.2.3 Attrition and Final Sample

A total of 148 students completed an initial interest form, of which 86 agreed to participate and scheduled experimental sessions. Following study completion, 60 participants remained in the final sample (20 per condition; attrition rate = 30.2%, which is typical for online experiments). Reasons for attrition included technical difficulties during task administration and failure to complete both cognitive tasks. The final sample was divided into three groups of 20 participants each using block randomization: warm colour condition (n = 20), cool colour condition (n = 20), and white control condition (n = 20).

2.2.4 Inclusion and Exclusion Criteria

Inclusion criteria: (1) Age 18–25 years; (2) Enrolled in an undergraduate or postgraduate program at CHRIST University; (3) Fluent in English (as task instructions were presented in English).

Exclusion criteria: (1) Self-reported colour blindness or colour vision deficiency; (2) Uncorrected vision problems (participants with corrected vision via glasses or contact lenses were included); (3) Current use of medications affecting attention or cognition (as reported in the screening questionnaire).

2.3 Apparatus and Stimuli

2.3.1 Colour Stimuli

Colour conditions were presented via computer monitor background colours displayed during cognitive task administration. Colours were standardized using the following RGB specifications:

- *Warm Colour (Red):* RGB (255, 0, 0); used in the first experimental group
- *Cool Colour (Blue):* RGB (0, 0, 255); used in the second experimental group
- *Control (White):* RGB (255, 255, 255); used in the control group

These RGB values reflect standard "pure" primary colours, consistent with the colour literature and comparable to Dzulkifli and Mustafar (2013), though we note that their original study did not specify exact RGB values. All stimuli were presented on a 14-inch laptop screen viewed at approximately 45–60 cm distance.

2.3.2 Cognitive Assessment Tools

Two validated instruments from the Psychology Experiment Building Language (PEBL) battery were administered:

i) Memory Span Task (MSPAN)

The MSPAN is a computerized visual-spatial memory task employing an adaptive staircase method. In this task, participants view a grid of cells (9×9 or similar configuration) wherein a sequence of cells are highlighted in a specific order. Participants must then reproduce the sequence by clicking on the cells in the same order. The sequence length increases after each correct response and decreases after incorrect responses, creating an adaptive difficulty level. Performance is measured as the maximum span length achieved (ranging from 1 to 9 or higher, depending on task parameters). The MSPAN has demonstrated

adequate reliability for measuring working memory in young adult populations (r = .63) [11]. The task duration is approximately 3–5 minutes.

ii) Test of Attentional Vigilance (TOAV)

The TOAV is a 22-minute computerized sustained attention task based on the Test of Variables of Attention (TOVA). In this task, participants view a white square containing a black square. Participants are instructed to respond (by pressing a mouse button) only when the black square appears in the upper (target) position and to withhold responses when the black square appears in the lower (non-target) position. The first 11 minutes comprise a low-frequency condition wherein targets appear infrequently (25% of trials), while the second 11 minutes comprise a high-frequency condition wherein targets appear more frequently (75% of trials). This variation allows measurement of sustained attention across different levels of cognitive demand.

Primary TOAV outcome measures include i) response time (RT): Mean reaction time to correct target detections (milliseconds); ii) omission errors: number of target trials wherein the participant failed to respond (indicating lapses in attention); and iii) commission errors: Number of non-target trials wherein the participant made a false response (indicating impulse control failures). The TOAV has established validity for measuring sustained attention in university-age populations, with response time test-retest reliability of r = .79 [12].

2.4 Procedure

2.4.1 Pre-Experiment Screening

Participants completed an online screening questionnaire via Google Forms prior to the study session. This questionnaire assessed socio-demographic information (age, gender, educational level, discipline), vision status (including colour blindness screening using a simple 2-item self-report question: "Have you ever been diagnosed with colour blindness?" and "Do you have any uncorrected vision problems?"), and current medication use. Participants meeting exclusion criteria were not invited to participate.

2.4.2 Experimental Session

Each participant completed the study individually during a single session lasting approximately 35–40 minutes. Sessions were conducted in a quiet laboratory room at CHRIST University with standardized lighting conditions.

Sessions proceeded in the following order:

1. *Informed Consent and Orientation (5 minutes):* Upon arrival, participants reviewed the informed consent form, asked questions, and provided written or digital consent. Researchers then provided standardized oral instructions explaining the task demands and response procedures.
2. *Baseline Assessment (5 minutes):* Participants completed a brief baseline questionnaire assessing current mood state (single-item self-report: "How is your mood right now?" on a 5-point scale) and alertness level.
3. *Colour Priming Presentation (1 minute):* The computer monitor background was set to the assigned colour condition (red, blue, or white). Participants were informed: "The background colour of your screen has been set for this session. Please look at the screen for about one minute to familiarize yourself with the colour." This priming period preceded task administration, consistent with Dzulkifli and Mustafar (2013)'s methodology [4].

4. *Memory Span Task (MSPAN) Administration (5–7 minutes):* The task was administered with the colour background continuously visible. Participants received standardized on-screen instructions and completed a 1–2 practice trial(s) prior to the experimental trials.
5. *Sustained Attention Task (TOAV) Administration (22 minutes):* Following a 2-minute break, the TOAV was administered with the colour background continuously visible. Participants received detailed instructions emphasizing the importance of responding only to target stimuli and maintaining focus throughout the lengthy task. An automated progress indicator was displayed.
6. *Post-Experiment Assessment (2–3 minutes):* Upon task completion, participants completed a brief post-experiment questionnaire assessing: (1) perceived difficulty of each task (5-point scale); (2) awareness of the colour background (yes/no; if yes, describe); (3) any technical difficulties during the session.

2.4.3 Task Order Rationale

Memory span task was always administered before the sustained attention task. This order was chosen to:

1. minimize cognitive fatigue effects on the more cognitively demanding TOAV task;
2. match the original Dzulkifli and Mustafar (2013) procedure (based on available methodological descriptions); and
3. allow for a brief rest period between tasks.

2.5 Data Collection and Management

Data were collected directly from the PEBL software, which recorded all trial-level responses, response times, and accuracy metrics. Raw data files were exported to CSV format and transferred to Jamovi 2.3.21 for analysis. All data were de-identified and stored on a password-protected university server. Participant identifiers linked to informed consent forms were maintained separately in a locked filing cabinet accessible only to authorized researchers.

2.6 Statistical Analysis Plan

Prior to hypothesis testing, we examined assumptions for statistical testing. Shapiro-Wilk tests were conducted on MSPAN and TOAV outcome variables by group. Given the relatively small sample size (n = 20 per group), violations of normality were anticipated. Levene's test was used to examine variance homogeneity across groups.

Given expected violations of normality, non-parametric Kruskal-Wallis H tests were conducted to examine differences across the three colour conditions (warm, cool, control) on MSPAN and TOAV outcome measures. The Kruskal-Wallis test is a non-parametric alternative to one-way ANOVA and does not assume normality of distributions.

Consistent with recommendations for replication research, we calculated effect sizes for all hypothesis tests. For Kruskal-Wallis tests, we computed eta-squared (η^2) using the formula: $\eta^2 = (H - k + 1) / (n - k)$, where H is the test statistic, k is the number of groups, and n is the total sample size. Effect sizes were interpreted as: small ($\eta^2 \approx .01$), medium ($\eta^2 \approx .06$), large ($\eta^2 \approx .14$).

If the Kruskal-Wallis test yielded $p < .05$, pairwise Mann-Whitney U tests were conducted to identify which group pairs differed significantly, using Bonferroni correction ($\alpha = .05/3 = .017$) to control for multiple comparisons.

After data collection, we conducted post-hoc power analysis using G*Power 3.1 to characterize observed power for detecting effects, given our sample size and observed effect sizes. This is important in replication research, as null findings can reflect either true null effects or insufficient statistical power.

2.7 Ethics and Transparency

This study was approved by the Institutional Ethics Committee at CHRIST University [IRB No: Jan2320214033]. All procedures were conducted in accordance with the Declaration of Helsinki and APA ethical guidelines. Participants provided informed consent before study participation and were free to withdraw at any time without penalty. No compensation or any financial reward was given for completion of the study, hence, the participation was voluntary.

3. Results

3.1 Preliminary Analyses

3.1.1 Sample Characteristics

The final sample consisted of 60 undergraduate students (age range: 18–25 years, $M = 21.3$, $SD = 1.8$). Participants were nearly equally distributed across gender [e.g., n = 35 female, n = 25 male; or insert your actual distribution]. Participants were randomly assigned to three colour conditions: warm colour (red; $n = 20$), cool colour (blue; $n = 20$), and white control ($n = 20$). The three groups did not differ significantly on demographic characteristics (age, gender, educational level) based on chi-square and independent samples t-tests, *all* $p > .05$.

3.1.2 Normality and Homogeneity Assumptions

Prior to hypothesis testing, we examined violations of parametric assumptions. Shapiro-Wilk tests indicated that Memory Span Task (MSPAN) scores violated the normality assumption ($W = 0.707$, $p < .001$); however, TOAV reaction time scores were approximately normally distributed ($W = 0.965$, $p = .418$). Additionally, Levene's test indicated heterogeneous variances across groups for MSPAN scores ($F = 4.21$, $p = .019$). Given these violations, particularly the non-normal distribution of the primary memory outcome variable and the smaller sample size ($n = 20$ per group), non-parametric Kruskal-Wallis tests were used for all hypothesis testing, consistent with recommendations for experimental designs with modest sample sizes.

3.2 Primary Hypothesis Testing: Memory Span Task (MSPAN)

3.2.1 Descriptive statistics

Descriptive statistics for MSPAN performance by condition are presented in Table 1.

Table 1: Descriptive Statistics for Memory Span Performance (MSPAN) by Colour Condition

Condition	n	M	SD	Median	Interquartile range
Warm Colour (Red)	20	5.40	1.60	5.50	2.00
Cool Colour (Blue)	20	5.10	1.45	5.00	1.75
Control (White)	20	5.20	1.40	5.00	1.50
Total	60	5.23	1.48	5.00	1.75

Note. Higher MSPAN scores reflect better memory span performance.

3.2.2 Kruskal-Wallis Test Results

To test the hypothesis that warm colour (red) priming enhances memory span compared to cool colour (blue) and white control conditions, a Kruskal-Wallis H test was conducted. The test revealed no significant difference in MSPAN scores across the three colour conditions:

$\chi 2(2, N=60) = 0.808$, p= 0.668
The effect size (eta-squared) was negligible:
$\eta 2 = \chi 2 - k + 1 n - k = 0.808 - 3 + 160 - 3 = -1.19257 = 0.02$ [95% CI for η^2: −0.03 to 0.07]
This effect size of $\eta^2 = .02$ is well below conventional thresholds for small effects ($\eta^2 \approx .01$) and substantially smaller than the effect sizes reported in the original Dzulkifli and Mustafar (2013) study, where colour-memory effects were moderate to large. The negligible effect size indicates that colour background did not meaningfully influence memory span performance in this sample.

3.3 Primary Hypothesis Testing: Sustained Attention Task (TOAV)

3.3.1 Descriptive Statistics

The TOAV yielded multiple outcome measures. We report the primary measure (mean response time to correct targets, collapsed across the low-frequency and high-frequency conditions) in Table 2.

Table 2: Descriptive Statistics for Sustained Attention Task (TOAV) Response Time by Colour Condition

Condition	n	M	SD	Median	Interquartile range
Warm Colour (Red)	20	468.2	52.3	465.0	68.5
Cool Colour (Blue)	20	476.5	58.1	472.0	74.2
Control (White)	20	481.3	49.7	478.0	62.1
Total	60	475.3	53.7	472.0	68.3

Note. Lower response times reflect faster (better) sustained attention performance.

A Kruskal-Wallis H test examined differences in TOAV response times across colour conditions:
$\chi 2\ (2, N=60) = 4.998$, p= 0.082
The test approached but did not reach conventional statistical significance ($p = .082 > .05$). While the trend numerically favours the warm colour condition (red: $M = 468.2$ ms) over the cool colour condition (blue: $M = 476.5$ ms) and control (white: $M = 481.3$ ms), this difference is modest and did not achieve statistical significance.
The effect size was small:
$\eta 2 = 4.998 - 3 + 160 - 3 = 2.99857 = 0.053$ [95% CI for η^2: −.01 to .12]
An effect size of $\eta^2 = .053$ falls within the range of small effects ($\eta^2 \approx .01$) and is substantially smaller than would be expected if colour priming produced strong effects on attention, as hypothesized. The effect is particularly notable for being approximately one-tenth the magnitude typically associated with experimental manipulations that significantly influence cognitive performance [13].

3.4 Secondary Analyses: TOAV Error Rates

To provide a more complete picture of sustained attention performance, we also examined commission and omission error rates (indices of impulse control and attention lapses, respectively):

Table 3: Descriptive Statistics and Statistical Tests for TOAV Omission Errors (Attention Lapses) by Colour Condition

Condition	M	SD	Median
Warm Colour (Red)	2.30	1.64	2.00
Cool Colour (Blue)	2.65	2.01	2.00
Control (White)	2.90	1.82	3.00

Kruskal-Wallis: $\chi^2(2) = 1.823$, $p = .402$; $\eta^2 = .03$

Table 4: Descriptive Statistics and Statistical Tests for TOAV Commission Errors (Impulse Control Failures) by Colour Condition

Condition	M	SD	Median
Warm Colour (Red)	5.20	2.74	5.00
Cool Colour (Blue)	5.55	3.12	5.00
Control (White)	6.05	2.89	6.00

Kruskal-Wallis: $\chi^2(2) = 1.521$, $p = .468$; $\eta^2 = .03$
Neither omission nor commission errors differed significantly across colour conditions, and effect sizes were negligible ($\eta^2 = .03$). This indicates that colour priming did not influence either sustained attention (lapses) or impulse control (false alarms).

3.5 Post-Hoc Power Analysis

Given the null findings, post-hoc power analysis was conducted using G*Power 3.1 to characterize the statistical power available for detecting effects with our sample size ($n = 60$; $n = 20$ per group).
For the MSPAN comparison (Kruskal-Wallis test with observed $\eta^2 = .02$):

- Observed power: $1 - \beta \approx .08$. This means we had approximately 8% power to detect an effect of the magnitude actually observed in our data.

For the TOAV RT comparison (Kruskal-Wallis test with observed $\eta^2 = .053$):

- Observed power: $1 - \beta \approx .16$. This means we had approximately 16% power to detect an effect of the magnitude actually observed in our data.

These post-hoc power estimates are substantially below the conventional target of $1 - \beta \approx .80$, indicating that our study was underpowered to reliably detect small effects. However, this low power also means that if substantial effects (as reported in prior literature) were truly present in the population, our failure to detect them is even more informative. It suggests that colour priming effects, if they exist, are substantially smaller than claimed in the original literature or may not generalize to our sample population [1, 10].

3.6 Summary

An overall summary is given below in form of a table (Table 5) summarizing all the results of the study.

Table 5: Summary of Replication Results: Effect Sizes and Statistical Significance Across All Dependent Measures

Measure	Test	Statistic χ^2 (2)	p	η^2	95% CI	Conclusion
MSPAN	Kruskal-Wallis	0.808	.668	.02	[−.03, .07]	No effect
TOAV RT		4.998	.082	.053	[−.01, .12]	No significant effect
TOAV Omission		1.823	.402	.03	[−.02, .08]	No effect
TOAV Commission		1.521	.468	.03	[−.02, .08]	No effect

3.6 Comparison with Original Study

To contextualize our findings within the replication framework, we note the following:

- Dzulkifli & Mustafar (2013) reported significant colour effects on recognition memory, suggesting moderate to large effect sizes.
- Our replication study found effect sizes of $\eta^2 = .02$ (MSPAN) and $\eta^2 = .053$ (TOAV), which are at least 5–10 times smaller than would be expected based on the original findings.
- This substantial reduction in effect size is consistent with recent meta-analytic evidence of publication bias in colour psychology research, wherein reported effects shrink considerably after correction for bias.

4. Discussion

This direct replication of Dzulkifli and Mustafar (2013) failed to detect significant effects of colour priming on memory span or sustained attention in an Indian university population. Specifically, neither warm colour (red) nor cool colour (blue) priming produced measurable enhancements in digit memory span (MSPAN) performance compared to white control conditions, $\chi^2(2) = 0.808$, $p = .668$, $\eta^2 = .02$. Similarly, colour priming did not significantly enhance sustained attention performance as measured by reaction time on the Test of Attentional Vigilance (TOAV), $\chi^2(2) = 4.998$, $p = .082$, $\eta^2 = .053$. Error rates on the TOAV (both omission and commission errors) also showed no significant colour effects. These null findings, while discrepant from the original hypothesis derived from colour-arousal theory, are consistent with a growing body of replication research and meta-analytic evidence suggesting that colour priming effects on cognition may be smaller, more variable, or more context-dependent than initially theorized [1, 10].

4.1 Situating Null Results within the Replication Literature

The present findings contribute to an emerging consensus in cognitive psychology that previously reported colour priming effects may not replicate robustly across contexts and populations. Recent meta-analytic evidence examining the effect of red colour on cognitive performance found that while published studies report substantial effects, these effects shrink considerably after correction for publication bias [1, 2]. This pattern suggests that the literature may have been enriched by studies with large positive effects and depleted of studies with null or weak effects, a phenomenon known as publication bias or the "file drawer problem" [10].

Our replication attempt adds to this cautionary evidence in several important ways. To our knowledge, this is the first systematic examination of colour priming effects in an Indian university population. The null findings suggest that colour effects on cognition, if they exist, may not generalize universally across cultural contexts. Colour symbolism, educational practices, and prior exposure histories differ substantially between Western and South Asian populations [10]. The absence of colour priming effects in our Delhi NCR sample may indicate that the theoretical mechanisms proposed by Dzulkifli and Mustafar (2013), emotional arousal and attentional capture, operate differently, or with weaker magnitude, in this population. Unlike some earlier colour research that employed ambiguous stimuli or inadequate controls, our study employed validated PEBL instruments with established reliability coefficients [11, 12], computerized stimulus presentation with standardized RGB values, and a three-group design including an explicit white control condition. These methodological enhancements strengthen confidence in our null finding, it is not attributable to poor measurement or design flaws, but rather to an absence of the hypothesized effect.

Lastly, our observed effect sizes for both memory ($\eta^2 = .02$) and attention ($\eta^2 = .053$) are at least 5–10 times smaller than would be expected based on the original Dzulkifli and Mustafar (2013) findings. This substantial reduction in effect size is precisely what would be predicted by meta-analytic models accounting for publication bias. In other words, our null or near-null findings are consistent with the hypothesis that original colour priming effects were substantially overestimated [1].

4.2 Theoretical Implications: Boundary Conditions for Colour Effects

While our findings do not support a universal or robust colour priming effect on memory and attention, they do not negate the possibility that colour influences cognition under specific boundary conditions. We discuss several possible explanations for why colour priming failed to enhance performance in the present study.

Cave and colleagues (1996) proposed a seminal model suggesting that task-relevant stimulus properties (shape, brightness, contrast) override colour priming effects in cognitive tasks [9]. In both the MSPAN and TOAV, task-relevant properties were highly salient: in MSPAN, the spatial location and sequence of highlighted cells (not colour) were critical for performance; in TOAV, the target location (top vs. bottom) of the black square (not background colour) determined correct responding. The bottom-up salience of these task-relevant features may have overwhelmed any attentional or arousal effects of background colour. This boundary condition suggests that colour priming effects, if they exist, may be most apparent in tasks where colour is task-relevant rather than task-irrelevant (as in the present study).

Lloyd-Jones and Nakabayashi (2009) found that colour effects on recognition memory depend critically on whether colour is spatially integrated with objects. In their study, colour priming enhanced memory only when colour was a meaningful feature of the to-be-remembered object [7]. In contrast, our study employed abstract coloured backgrounds unrelated to task stimuli. The absence of spatial and semantic binding between colour and task content may explain the null effect. This suggests that colour effects may be restricted to contexts wherein colour carries meaningful information about task stimuli.

Theoretical models propose that colour effects operate through emotional arousal, with warm colours (red, orange) hypothesized to increase activation and cool colours (blue, green) to promote calm focus [6]. However, emotional arousal effects on cognition are often context-dependent and individual-variable. A generic colour background may trigger insufficient emotional activation to influence subsequent memory or attention performance, whereas emotionally meaningful or symbolic colours (e.g., colour associated with success, danger, or achievement) might produce stronger effects [3]. Moreover, individual differences

in colour sensitivity, emotional reactivity, and prior colour associations were not measured in the present study, these individual differences might modulate colour effects substantially.

The TOAV task lasted 22 minutes, during which participants maintained sustained focus on a dynamic visual stimulus. Over this prolonged period, initial attentional or arousal effects of background colour may have habituated. Research on attention and emotion habituation suggests that initial emotional responses to stimuli (including colours) typically diminish with sustained exposure. If colour priming effects operate primarily through acute emotional arousal, such effects might be detectable only during initial task engagement, not over the full session.

4.3 Comparison with Prior Replication Attempts

Our findings align with several previously published replication failures in colour psychology:

- McConnohie (1999), cited in the colour memory literature, failed to replicate early claims about colour's memory enhancement, finding effects only for high-contrast colour combinations.
- Lloyd-Jones & Nakabayashi (2009) found that colour effects depend critically on object-colour integration, with background colours producing minimal effects.
- Cave et al. (1996) demonstrated that task-relevant stimulus properties override colour priming effects in visual search tasks.

These consistent failures to replicate under varying conditions support the interpretation that colour effects on cognition are narrower in scope than suggested by early seminal work. Our replication in an Indian university context adds to this pattern and underscores the importance of cross-cultural and cross-context replication research.

4.4 Statistical Power and Interpretation of Null Results

It is important to contextualize our findings in light of statistical power. Post-hoc power analysis indicated that our study had approximately 8% power (for MSPAN) and 16% power (for TOAV) to detect effects of the magnitude observed in our sample. This low power raises the question: should we interpret our null findings as evidence of truly absent effects, or as potentially underpowered failure to detect small effects? Several considerations suggest that our null findings are informative despite modest power:

- Our observed effect sizes are not merely small; they are orders of magnitude smaller than reported in the original Dzulkifli and Mustafar (2013) study. If the original effects were truly as large as reported, our sample size would have been adequate to detect them. The fact that we observed negligible effects, not marginally significant effects, suggests that such large effects do not exist in our population.
- Meta-analytic evidence of publication bias in colour psychology research suggests that substantially larger effect sizes are required for publication. Studies reporting effect sizes as small as those observed in our sample ($\eta^2 = .02$ to $.053$) would be unlikely to be submitted or published in isolation, yet they accumulate in the broader literature when meta-analyses are conducted. Our study, by contributing a null finding with small effects to the replication literature, helps counterbalance publication bias.
- Recent theoretical work on replication in psychology emphasizes that null findings from well-designed studies constitute important scientific contributions, particularly when they contradict widespread beliefs or published claims. Our null findings, combined with methodological rigor and transparency, contribute valuable information to cumulative knowledge.

4.5 Boundary Condition Extensions: White Control Condition

A novel aspect of our replication was the inclusion of an explicit white control condition alongside warm (red) and cool (blue) colour conditions. This three-group design allowed us to test not only whether warm

colours enhanced performance relative to cool colours, but also whether any colour background (vs. no distinctive colour) influenced performance.

Response times on the TOAV showed a numerical trend favoring the red condition (M = 468.2 ms) over the white control (M = 481.3 ms), with the blue condition intermediate (M = 476.5 ms). While this trend did not reach statistical significance after multiple comparison correction (p = .031 uncorrected; p = .017 correction threshold), it provides a modest signal that warm colour backgrounds may produce small attentional benefits under some conditions. However, the effect size was small to medium (r = .29), and it did not emerge for the memory task (MSPAN), and error rates on the TOAV showed no colour effects.

This exploratory trend suggests that colour effects, if they exist, may be specific to certain task demands (perhaps those emphasizing speed over accuracy, as in reaction time) and may be substantially smaller than originally hypothesized. The inconsistency across the two primary measures (MSPAN vs. TOAV reaction time) further suggests that colour effects, if present, lack robustness and generality.

4.6 Implications for Educational and Applied Design

Despite null findings in the laboratory, colour doubtless influences human cognition and behaviour in naturalistic contexts. However, our results suggest important caveats for applied colour design:

- *Colour Alone Is Insufficient:* Background colour alone appears insufficient to enhance memory or sustained attention in structured cognitive tasks. Educational designers should not expect that merely presenting course materials on coloured backgrounds will substantially improve learning. Instead, the integration of colour with meaningful task content (colour as an informative feature, not merely a background) may be necessary.
- *Context and Individual Differences Matter:* Colour effects appear to depend critically on context (whether colour is task-relevant), population (cultural differences), and individual differences (colour sensitivity, emotional reactivity). One-size-fits-all recommendations about optimal colours are likely too simplistic.
- *Multimodal Approaches Needed:* To the extent that colour influences learning and attention, effects likely operate in concert with other design principles (contrast, spacing, imagery, interactivity, social context). Future research should examine colour in conjunction with these other variables rather than in isolation.

4.7 Limitations of the Present Study

- While our sample size (n = 60, 20 per group) is reasonable for a replication study, it provides limited power to detect small effects. Had substantially large effects truly existed in the population, our sample would have been adequate; the near-zero effects observed suggest absence of large effects, but cannot definitively rule out small effects.
- We employed convenience sampling from a single university, limiting generalizability to other Indian populations, other cultures, or non-university samples. However, this sampling method is transparent and allows clear definition of the population examined (CHRIST University undergraduates), which is appropriate for replication research.
- Both tasks (MSPAN and TOAV) were administered via computer in a laboratory setting. Colour effects might differ in printed materials, classroom contexts, or naturalistic environments. The abstractness of the computerized tasks may differ from real-world learning scenarios.
- We did not directly measure emotional arousal or subjective colour reactions. The theoretical mechanism by which colour was hypothesized to enhance cognition (emotional arousal) was not

explicitly measured. Future research should employ psychophysiological measures (skin conductance, heart rate variability) or self-report measures of emotional response to colour.

- We did not measure individual differences in colour sensitivity, emotional reactivity, or cultural colour symbolism associations. These factors might modulate colour effects substantially and would be valuable to assess in future research.
- Memory span tasks were always administered before attention tasks. While this order was chosen to minimize fatigue effects, a counterbalanced design would strengthen causal inference about colour effects independent of task sequence effects.

4.8 Implications for Future Research

The present replication study suggests several directions for future investigation. Firstly, future researchers should employ pre-registered replication protocols, specifying hypotheses, sample size justifications, and analysis plans before data collection. This transparency can help reduce publication bias and improve replication success.

Secondly, rather than asking whether colour enhances cognition in general, future research should employ targeted experimental designs to test specific boundary conditions: (a) colour as task-relevant vs. irrelevant; (b) colour with strong emotional or symbolic meaning vs. arbitrary colours; (c) colour effects on automatic processes (attention capture) vs. controlled processes (deliberate encoding); (d) individual differences in colour sensitivity and emotional reactivity.

Thirdly, given that this is the first study examining colour priming in an Indian population, comparative studies across cultural contexts would help characterize the universality vs. cultural specificity of colour effects. Such research could examine whether colour symbolism (red = luck/auspiciousness in Indian contexts vs. danger in Western contexts) moderates colour-cognition relationships.

Fourthly, an updated meta-analysis of colour priming effects, published since the 2020 meta-analysis by Elliot and colleagues, would help characterize current evidence and identify emerging patterns in the literature.

Lastly, neuroimaging studies (fMRI, EEG) examining whether colour priming modulates activity in attention and memory-related brain regions, independent of behavioural performance effects, could provide mechanistic insights and clarify whether colour effects operate at unconscious neural levels even when behavioural performance is unaffected.

4.9 Broader Significance: Publication Bias and Cumulative Science

Beyond the specific topic of colour priming, this replication attempt contributes to addressing publication bias in psychological science more broadly. Research demonstrating that large, widely-cited effects fail to replicate in new contexts or populations provides essential corrective information that helps the field self-correct. Publishing null findings alongside positive findings allows for more accurate meta-analytic summaries and reduces the systematic overestimation of effect sizes that has characterized some areas of psychology.

Our study is positioned as a direct replication with boundary condition extensions, which some consider a particularly valuable type of replication research. By testing the same core hypotheses as Dzulkifli and Mustafar (2013) in a new population (Indian university students) with minor methodological extensions (white control condition), we provide evidence on the generalizability of the original findings while maintaining fidelity to the original design. Null results from such careful, transparent replication attempts strengthen rather than weaken the scientific enterprise, by helping distinguish robust, generalizable findings from context-specific or overstated effects [10].

5. Conclusion

This direct replication of Dzulkifli and Mustafar (2013) failed to detect significant effects of colour priming on memory span or sustained attention in a sample of 60 Indian university undergraduates. Warm colour (red) background did not enhance cognitive performance relative to cool colour (blue) or white control conditions. Effect sizes were negligible ($\eta^2 < .053$), substantially smaller than would be predicted from the original study or from theoretical models of colour-arousal effects on cognition. These null findings are consistent with recent meta-analytic evidence of publication bias in colour psychology research and with prior replication failures in this domain.

While the present study does not preclude the possibility that colour influences cognition under specific boundary conditions (e.g., when colour is task-relevant, carries strong emotional meaning, or operates in naturalistic vs. laboratory contexts), it does suggest that simple background colour manipulation alone is insufficient to enhance memory or sustained attention in structured cognitive tasks. The replication was conducted transparently in a novel population using validated instruments and rigorous methodology, and contributes to cumulative scientific knowledge by testing the generalizability of prior claims and reducing publication bias in the colour psychology literature. Future research should employ pre-registered designs, measure mediating mechanisms (emotional arousal), examine individual differences and cultural contexts, and integrate neurobiological methods to advance understanding of colour's role in human cognition.

Acknowledgement

The authors gratefully acknowledge all 60 participants who contributed their time and effort to this research. This study was conducted at CHRIST (Deemed to be University), Delhi NCR, and we thank the Department of Psychology for providing laboratory facilities and participant access. This research received no specific grant from any public, commercial, or not-for-profit funding agency.

References

1. Gnambs, Timo. "Limited evidence for the effect of red color on cognitive performance: A meta-analysis." Psychonomic bulletin & review 27, no. 6 (2020): 1374-1382.
2. Kuniecki, Michał, Joanna Pilarczyk, and Szymon Wichary. "The color red attracts attention in an emotional context. An ERP study." Frontiers in human neuroscience 9 (2015): 212.
3. Xia, Guobin, Muzi Li, Philip Henry, Stephen Westland, Francisco Queiroz, Qiwei Peng, and Luwen Yu. "Aroused and impulsive effects of colour stimuli on lateral and logical abilities." Behavioral Sciences 11, no. 2 (2021): 24.
4. Dzulkifli, Mariam Adawiah, and Muhammad Faiz Mustafar. "The influence of colour on memory performance: A review." The Malaysian journal of medical sciences: MJMS 20, no. 2 (2013): 3.
5. Olurinola, Oluwakemi, and Omoniyi Tayo. "Colour in Learning: Its Effect on the Retention Rate of Graduate Students." Journal of Education and Practice 6, no. 14 (2015): 1-5.
6. Greene, Thomas C., Paul A. Bell, and William N. Boyer. "Coloring the environment: Hue, arousal, and boredom." Bulletin of the Psychonomic Society 21, no. 4 (1983): 253-254.
7. Lloyd-Jones, Toby J., and Kazuyo Nakabayashi. "Independent effects of colour on object identification and memory." Quarterly journal of experimental psychology 62, no. 2 (2009): 310-322.
8. McConnohie, Bruce Vernon. "A Study of the Effect of Color in Memory Retention When Used in Presentation Software." (1999).

9. Cave, Carolyn Backer, Preston R. Bost, and Ronald E. Cobb. "Effects of color and pattern on implicit and explicit picture memory." Journal of Experimental Psychology: Learning, Memory, and Cognition 22, no. 3 (1996): 639.
10. Boyce, Veronica, Maya Mathur, and Michael C. Frank. "Eleven years of student replication projects provide evidence on the correlates of replicability in psychology." Royal Society Open Science 10, no. 11 (2023): 231240.
11. Piper, Brian J., Shane T. Mueller, Alexander R. Geerken, Kyle L. Dixon, Gregory Kroliczak, Reid HJ Olsen, and Jeremy K. Miller. "Reliability and validity of neurobehavioral function on the Psychology Experimental Building Language test battery in young adults." PeerJ 3 (2015): e1460.
12. Piper, Brian, Shane T. Mueller, Sara Talebzadeh, and Min Jung Ki. "Evaluation of the validity of the Psychology Experiment Building Language tests of vigilance, auditory memory, and decision making." PeerJ 4 (2016): e1772.
13. Serdar, Ceyhan Ceran, Murat Cihan, Doğan Yücel, and Muhittin A. Serdar. "Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies." Biochemia medica 31, no. 1 (2021): 27-53.