Reverse Stroop Effect: The Plasticity of Interference Contingent Upon Mode of Output

Sarah L. Sindeband

Florida Atlantic University

PSY 3213: Research Methods in Psychology

Dr. Terrence Barnhardt

November 30th, 2021

Abstract

The purpose of this study was to support the idea that participants performing a Reverse Stroop Task would encounter more interference when the task was to name a word, rather than the ink color, of an incongruent stimulus. Forty-four undergraduate students enrolled in a Research Methods course at Florida Atlantic University participated in the study. The modified (Reverse) Stroop paradigm (Stroop, 1992) was analyzed with a 2-way repeated measures ANOVA. The 2x2 factorial design had two independent variables being congruence and task, and two dependent variables being accuracy and reaction time. The results supported the hypothesis suggesting the plasticity of the interference in performing Stroop tasks is dependent on the mode of output of the task. This was demonstrated by using ink color as the mode of output, thus resulting in less interference in the ink task in comparison to the word task.

Keywords: cognition, Stroop Effect, Reverse Stroop Effect, plasticity of perception

Reverse Stroop Effect: The Plasticity of Interference Contingent Upon Mode of Output

The mechanism of creating a conflict between fast and relatively slower mental processes behind the Stroop Effect (Stroop, 1992) has been explored in many different experiments since the original study in 1935. The interference being manipulated in the subsequent research investigates possible moderating and mediating variables intrinsic to the parallel processes involved in performing the variations of the Stroop task. The Reverse Stroop effect in the present study aims to produce an opposite effect by manipulating the mode of output in the task to demonstrate the plasticity of the automatic nature of the cognitive processes involved.

The notion that the stimulus input of the incongruent ink color might have caused the interference in the ink task has been challenged. This was tested by comparing whether the Stroop Effect is caused by interference from selective attention of the encoding process or if selective attention hinders the process following the memory recall process for the correct output option (Hintzman et al., 1972). Hintzman et al. suggested that if the Stroop Effect was caused by the mechanism of word processing being automatic, and the ink color alone created interference, the word matching its ink color would have same level of interference as the word that was irrelevant to the color of the ink. The results of the experiment showed more interference in the condition with a made-up word than the word with the congruent ink color (Hintzman et al., 1972) which supported the idea that the response competition created the interference in the Stroop Effect through the process of the congruent stimulus priming the correct response. From this, one can infer that while completing the tasks in the Stroop experiment (Hintzman et al., 1972), selective attention is influencing the thought process after the stimulus is recalled from short-term memory. The decrease in interference for the congruent condition in contrast to the made-up word (Hintzman et al., 1972) supports the claim that the congruent color of the font accelerated the process of obtaining the right answer. This leads to the plasticity with the automatic nature of the comprehension of the stimulus being dependent on the mode of output.

In the experiments from the original Stroop article (Stroop, 1992), the participants identified the correct answers with a verbal response. This created an asymmetrical decrease in interference for the incongruent word task in comparison to the ink task. To counterbalance the practice effects of verbally reciting color words, (Blais & Besner, 2007) a more neutral form of indication (pointing to a color) was used for response indication. This method was tested on both Stroop and Reverse Stroop tasks. In the Stroop tasks (Stroop, 1992), the translation would occur when the goal was to indicate a font color, which was seeking a word output. The Reverse Stroop tasks (Blais & Besner, 2007) would involve translation when the output was a color, and the participant would need to match the written word, in an incongruent font color, to the color outputs. When the translational aspect of the verbal response was removed, the data still shows a Reverse Stroop Effect when pointing out the color words with incongruent ink colors (Blais & Besner, 2007). Therefore, removing the asymmetrical nature of the interference based on the association between word output options and the verbal pathway of indicating the correct answer.

A series of experiments manipulating the modes of input and output in Stroop and Reverse Stroop tasks (Sobel et al., 2020) compared the visuospatial and verbal associations involved. One experiment in the series (Sobel et al., 2020) increased interference in a regular Stroop task by making the words more difficult to read. This slowed down the process of reading the word, which allowed the color of the ink to induce that increase in interference, creating a Reverse Stroop Effect (Stroop, 1992). The association between the verbal pathway of output facilitating the response in a regular Stroop task demonstrated the asymmetrical interference caused by the association (Sobel et al., 2020). The opposite was observed when localizing indication methods were employed upon

Stroop and Reverse Stroop tasks (Sobel et al., 2020). The visuospatial association facilitated a decrease in the interference of the Reverse Stroop Effect when pointing out colors of incongruent word conditions of the Reverse Stroop task (Sobel et al., 2020). These findings supported the hypothesis that this neutral pointing method of indication would have no more practice effects for word output options than the color output mode in comparison to the verbal indication method.

Another study investigating the Reverse Stroop Effect also compared the Stroop paradigm (Ikeda et al., 2010). There was less interference for the Reverse Stroop (Ikeda et al., 2010) tasks for both modes of feedback (nonverbal and verbal) and more interference for the Stroop tasks (Stroop, 1992) in the verbal feedback than the nonverbal feedback. The study speculated that congruence has a greater effect on verbal stimuli than sensory stimuli (Ikeda et al., 2010). This would account for the greater interference in the Stroop tasks than the Reverse Stroop tasks (Ikeda et al., 2010). This study, like the present study, addressed the threat to construct validity that unsystematic representation of stimulus.

One can speculate that the interference caused by the Stroop Effect and Reverse Stroop Effect is due to the variance in the mental steps required to perform the tasks under each condition. The present study investigated the process of inducing the Reverse Stroop Effect by manipulating the mode of output for word matching and ink matching tasks from the original Stroop experiments. The within-subjects design featured two independent variables: task and congruence. The tasks the participants performed were three adaptations of the Stroop task (Stroop, 1992). Congruence was quantified by whether the written word stimulus presented in the ink and word tasks matched the font color it was presented in (congruent) or not (incongruent). The dependent variables were the reaction time for each response and the accuracy of each response. Response time was measured in seconds. Accuracy was measured by how many tasks the participant responded correctly and presented as a percentage of correct answers. The main effects of task and congruence were expected to yield the most accurate and quickest responses in the ink task and the congruent condition. The interaction effect, regarding task and congruence was hypothesized to have the most accurate and quickest responses for the congruent trials of the ink task.

Method

Participants

The participants (*N* = 44) were undergraduates (36 females, 8 males) enrolled in the lecture of a Research Methods Course at Florida Atlantic University. The average age was 22.39 (*SD* = 4.48, range: 18-49). The author of this paper (S.L.S.) was one of the participants in this study. Participants answered the demographics survey stating they were a Non-Native English speaker (*n* = 12) or a Native English speaker (*n* = 32). All participants also reported being right-handed. All participants enrolled in the class completed all the Reverse Stroop tasks in the experiment, there were no exclusions. The participants did not receive any compensation or incentive in return for participation in the in-class exercise. The study was approved by Florida Atlantic University's Institutional Review Board and the treatment of the participants was compliant with the "Ethical Principles of Psychologists and Code of Conduct" (American Psychological Association, 2017).

Materials

The Reverse Stroop tasks (Stroop, 1992) and demographics survey were performed on desktop computers in the designated classroom, using EPrime 3 Software, in English. The three colors of the hashtags used in the tasks were red, green, and blue. The participants pressed the corresponding key on the number pad of a standard keyboard with their right hand to submit their answers. The standardized instructions for all the tasks were provided by the instructor.

Design

The experiment featured a 2 x 2 within-subjects design. The dependent variables measured were accuracy (i.e., correct vs. incorrect) and response time (RT). The independent variables were the task (i.e., word task vs. ink task) and congruency (i.e., congruent vs. incongruent). The input for all three tasks was displayed in the center of the screen for each participant. All the participants completed the three tasks in the same order, during their scheduled class time. The covariates were sex of the participants and language (Native English speaker or Non-Native English speaker).

Procedure

The study took place in a classroom during the scheduled class time in the second week of the semester. The teaching assistant for the designated lab section of each course (3 lab course sections) gave a set of standardized instructions for each task, before the participants began. The participants were told to focus on accuracy (90% correct or above) rather than speed. Each participant completed a total of forty-five trials in each of the three tasks, as five sets of nine tasks. Order was not controlled for because the participants performed all the tasks in the same order. The setting was controlled for by having the participants in each lab section completed all the trials on the same desktop computer, together, in the same classroom.

Before each stimulus appeared on the screen, a fixation cross was presented for one second before the stimulus. The stimulus would appear for five seconds or until the participant responded on the keyboard with the number corresponding to the color. The output options for all three tasks were the same. They consisted of three sets of colored hashtags (red, green, or blue) and a number (1-####) that corresponded to the color (1 was red, 2 was green, 3 was blue) presented at the bottom of the screen. After the response was submitted, feedback was presented on the screen. This feedback included whether their response for that trial was correct, the overall accuracy, and the response time for that trial in seconds. The feedback screen was displayed for one and a half seconds.

The first task the participants completed was the practice task. The stimulus was a set of hashtags which would appear in one of three font colors (red, green, or blue) per trial. The practice task served as a control measure because there was no control group. The practice task required the participants to match the color of the font in which the hashtags were displayed to the color output options displayed below. The second task the participants completed was the ink task. This task required the participants to match the color of the font in which the word stimulus was displayed to the color output options at the bottom of the screen. The ink task had a congruent or incongruent condition in which the ink color could match the written word, or the written word would be incongruent with the ink color. The third task the participants completed was the word task. The word task required the participants to match the written word (red, green, or blue) to the color output options displayed on the bottom of the screen. The stimulus could be presented in a congruent condition or an incongruent condition. The congruency depended on if the font color (red, green, or blue) matched the written word displayed (red, green, or blue).

Each color word stimulus was presented equally (three times) in each set of nine tasks, therefore had the same chance of being displayed in each color. This created a congruency (i.e., the word “red” displayed in the red font) condition of the stimulus for the word task and ink task one third of the time. The incongruent (i.e., the word “red” displayed in blue font) condition occurred two thirds of the time. Completing all forty-five trails took the participants approximately twenty minutes in the single session, on the same computers. Then the teaching assistant debriefed the participants, which included the logic behind the Stroop Effect (Stroop, 1992) and the expected Reverse Stroop Effect.

Data & Analysis

After all the data were collected the group averages for response time and accuracy were analyzed. There were no data missing from the results gathered. The professor of the Research Methods class analyzed the data by running multiple two-way repeated measures ANOVA and multiple paired samples *t*-tests with (SPSS) software.

Results

In the present study of the Reverse Stroop Effect, it was hypothesized that the level accuracy would be higher on the ink tasks in the congruent trials. A two-way repeated measures analysis of variance (ANOVA) was used to tell if there was a significant difference in the data for the main effect attributed to task (word or ink) on the accuracy of the responses (See Table 1). The ANOVA indicated the main effect of task (word or ink) on accuracy was significant $F(1, 43) = 17.30$, $MSE = .002$, $p < .001$, and the effect size was small ($\eta_p^2 = .29$). The mean of the accuracy for the ink task ($M = .97$, $SE = .01$) yielded a significantly higher result than the word condition ($M = .95$, $SE = .01$). The ANOVA used to test the main effect attributed to congruence (congruent or incongruent) on the accuracy of the responses also produced a significant result; $F(1, 43) = 46.77$, $MSE = .002$, $p < .001$, and the effect size was large ($\eta_p^2 = .51$). The mean of accuracy for the congruent trials ($M = .98$, $SE = .004$) yielded a significantly higher result than the incongruent condition ($M = .93$, $SE = .01$). The results, shown in Figure 1, of the ANOVA that tested the interaction between task and congruence on accuracy were significant $F(1, 43) = 14.28$, $MSE = .003$, $p < .001$, and the effect size was small ($\eta_p^2 = .25$). The mean of accuracy for the ink tasks with congruent conditions ($M = .99$, $SE = .01$) was significantly higher than those of the incongruent conditions of the ink task ($M = .91$, $SE = .01$), the congruent conditions of the word task ($M = .98$, $SE = .01$), and the incongruent conditions of the word task ($M = .96$, $SE = .01$).

Two paired samples *t*-tests were used to analyze the simple effects of the two-way interactions after the ANOVA tests established the significant results to support the hypotheses for accuracy. The effect of accuracy among congruence was expected to be significant for both ink and word task, but greater for the word task. The paired samples t-test for accuracy was significant for the word task among the levels of congruence $t(43) = 6.10$, $p < .001$, and the effect size was large ($d = .92$). The paired samples t-test for accuracy was also significant for the ink task among levels of congruence $t(43) = 2.32$, $p = .03$, and the effect size was medium ($d = .35$). The effect of congruence on accuracy was greater for the word task ($M = .08$, $SD = .09$) than it was on the ink task ($M = .02$, $SD = .05$). The analyses of the main effects of task and congruence, the interaction effect of task and congruence on accuracy, and the simple effects of the two-way interaction of congruence among the levels of task of accuracy confirmed the hypotheses of the study. Therefore, the null hypotheses were rejected with a 95% alpha level.

Response time was expected to decrease on the ink tasks in the congruent trials. A two-way repeated measures analysis of variance (ANOVA) was used to determine significance in the difference in the data for the main effect attributed to task (word or ink) on the reaction time of the responses. The ANOVA indicated the main effect of task (word or ink) on reaction time was significant $F(1, 43) = 38.17$, $MSE = 10403.23$, $p < .001$, and the effect size was medium (η_p^2 =

.47). The mean of the reaction time for the word task (*M* = 716.29, *SE* = 26.08) yielded a significantly slower reaction time than the ink condition (*M* = 621.30, *SE* = 29.16). The ANOVA used to test the main effect attributed to congruence (congruent or incongruent) on the reaction time also produced a significant result; $F(1, 43) = 80.31$, $MSE = 16916.60$, $p < .001$,and the effect size was large ($\eta_p^2 = .65$). The mean reaction time for the congruent trials (*M* = 581.04, *SE* = 19.54) yielded a significantly quicker response time than the incongruent condition (*M* = 756.56, *SE* = 34.96). The results of the ANOVA that tested the interaction between task and congruence on response time were significant $F(1, 43) = 35.59$, $MSE = 4252.75$, $p < .001$, and the effect size was medium ($\eta_p^2 = .45$). The mean of response time for the ink tasks with congruent conditions (*M* = 599.21, *SE* = 18.07) was significantly faster than those of the incongruent conditions of the ink task (*M* = 833.38, *SE* = 37.10), the congruent conditions of the word task (*M* = 562.86, *SE* = 23.46), and the incongruent conditions of the word task (*M* = 697.73, *SE* = 35.90). Even though the results for the interaction were significant, there was no interaction between task and congruence on response time.

Two paired samples *t*-tests were used to analyze the simple effects of the two-way interactions after the ANOVA tests established the significant results to support the hypotheses for response time. The effect of response time among congruence was expected to be significant for both ink and word task, but greater for the word task. The paired samples *t*-test for reaction time was significant for the word task among the levels of congruence $t(43) = -8.95$, $p < .001$, and the effect size was large ($d = -1.35$). The paired samples *t*-test for response time was also significant for the ink task among levels of congruence $t(43) = -7.02$, $p < .001$, and the effect size was large ($d = -1.35$). The effect of congruence on reaction time was greater for the word task (*M* = -234.18, *SD* = 173.62) than it was on the ink task (*M* = -116.87, *SD* = .110.43). The analyses of the main effects of task and congruence, the interaction effect of task and congruence on accuracy, and the simple effects of the two-way interaction of congruence among the levels of task of response time confirmed the hypotheses of the study. Therefore, the null hypotheses were rejected with a 95% alpha level.

Discussion

The present study investigated whether the Reverse Stroop Effect (RSE) would produce the opposite results of the Stroop Effect. The purpose of this study was to support the idea that the participants would encounter more interference when the task was to name the ink color, rather than the name of the word, of an incongruent stimulus. Accuracy, regarding the main effects of congruence and task, was hypothesized to have the best scores in the congruent trials and on the ink task. The prediction for the interaction effect of task and congruence on accuracy was that the congruent trials of the ink task would yield the highest rate of accuracy. Reaction time, regarding the main effects of congruence and task, was hypothesized to have the shortest reaction time for the congruent trials and on the ink task. The prediction for the interaction effect of task and congruence on reaction time was that the congruent trials of the ink task would yield the shortest reaction times. These results would indicate the opposite of what is expected from a traditional Stroop experiment with the mode of output being a word to match the stimuli to instead of the ink color. The data supported the hypotheses that there was greater interference (longer reaction time and decreased accuracy) in the incongruent condition of the word task than the incongruent condition ink task when the goal was to match a color output (RSE) rather than a word output as seen in a traditional Stroop experiment. There was a significant difference in the means for the main effects of congruence and task on the reaction time and accuracy. The analysis of the interaction effects for the congruence and task showed a significant difference in both the reaction time and the accuracy.

When the Stroop Effect was originally investigated in 1932 (Stroop, 1992), participants were given words (names of colors) to match with congruent or incongruent font colors. In the word task, the participants would match the word written with their word output options and respond verbally in the ink task, the participants would match the color of the ink to the word output options and respond verbally (Stroop, 1992). Results from both the ink task and the word task showed an increase in interference form the incongruent stimulus compared to the congruent stimulus (Stroop, 1992). There was an increase (74.3%) in response time for naming the words of the congruent condition compared to naming the words in a conflicting font color (Stroop, 1992). This supported the hypothesis that the amount of interference was greater in the ink task than it was in the word task. Even when the ink task was practiced, the decrease in interference was short-lived (Stroop, 1992). The mode of output being only word responses, which were only verbally expressed, created an asymmetrical interference between the two modes (color or word) of stimulus (Stroop, 1992). One could speculate that increased interference (longer reaction time and decreased accuracy) in the Reverse Stroop experiment from the present study was induced in the word task due to the incongruent font color. In the present study when the output was a color, instead of a word, the stimulus did not need to be translated to match the color output options, therefore, encountered less interference (longer reaction time and decreased accuracy) in the congruent condition of the ink task. This demonstrates the plasticity in which the process is classified as more automatic depending on the goals or output mode of the task.

A study comparing the Stroop Effect interference (SI) and the Reverse Stroop Effect interference (RI) by measuring brain activity found that there are two separate processes facilitating that interference (Song & Hakoda, 2015). This series of experiments measured activity in different brain areas, rather than measuring behavioral responses such as those in the present study. The differences in the activity can be due to the different mode of output for response that the tasks required (Song & Hakoda, 2015). The RI was observed when the response required a physical response rather than a verbal one (Song & Hakoda, 2015). One can speculate that this could be due to the association between colors being identified more so in real life as visuospatial elements than by oral responses.

Validity, Limitations & Future Directions

The present study included a baseline in the original data that was not included in the final analyses. This baseline condition of the ink task showed less interference than the congruent condition for measures of accuracy and reaction time. This demonstrates that even though the stimulus was congruent, the association of the word still caused more interference than the neutral condition of the baseline of colored hashtags as input. The baseline condition for the word task showed more interference than the congruent condition for both accuracy and response time. This demonstrated the congruent condition of the word matching the font color facilitated the process of responding correctly. The baseline condition for the word task could technically be counted as incongruent because the ink was black, therefore causing more interference than the congruent condition. This threat to construct validity was not a parallel manipulation of the baseline in comparison to the neutral baseline used for the ink task.

Furthermore, the sample for this experiment only included right-handed participants. A more diverse sample for future experiments with some left-handed participants might yield different results and strengthen the construct validity. The participants for the current study were split into three separate lab sections, at different times, in the same room, with different instructors providing the instructions for them. The instructions were standardized to minimize interpretation. All the participants were told the speed and accuracy were important in all of the tasks, but to maintain at least 90% accuracy. This did not seem to affect response time because all analyzed response times were still significant. Internal validity for future studies could be addressed by comparing the Stroop Effect and the Reverse Stroop effect in a within-groups design.

Conclusion

The Reverse Stroop paradigm in the current study supported the hypotheses that it would produce the opposite results of the original Stroop paradigm (Stroop, 1992). Instead of the congruence having a greater effect on the ink task, the results showed the word task yielding a greater effect of congruence. The results of the present study suggest that the automatic nature of the processing required to complete Stroop and Reverse Stroop tasks is dependent on the mode of output that the task requires. When the mode of output does not match the stimulus, such as in the word task of the RSE, there is a greater effect of congruence than in the ink task. This was also observed in the Stroop Effect (Stroop, 1992) for the ink task because the input of ink color did not match the mode of output which was a word.

References

American Psychological Association. (2017). Ethical Principles of Psychologists and Code of Conduct. Retrieved from https://www.apa.org/ethics/code#

Blais, C., & Besner, D. (2007). A reverse Stroop effect without translation or reading difficulty. *Psychonomic bulletin & review*, *14*(3), 466-469. https://doi.org/10.3758/BF03194090

Hintzman, D. L., Carre, F. A., Eskridge, V. L., Owens, A. M., Shaff, S. S., & Sparks, M. E. (1972). "Stroop" effect: input or output phenomenon? *Journal of experimental psychology*, *95*(2), 458-459. https://doi.org/10.1037/h0033644

Ikeda, Y., Hirata, S., Okuzumi, H., & Kokubun, M. (2010). Features of Stroop and reverse-Stroop interference: Analysis by response modality and evaluation. *Perceptual and motor skills*, *110*(2), 654-660.

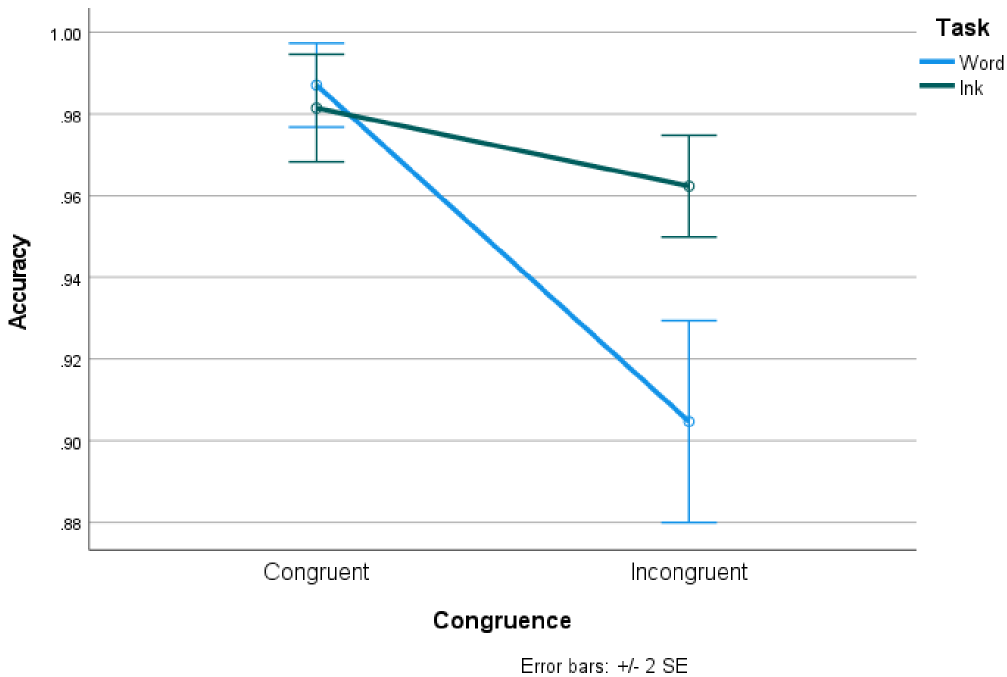
Sobel, K. V., Puri, A. M., & York, A. K. (2020). Visual search inverts the classic Stroop asymmetry. *Acta psychologica*, *205*, 103054. https://doi.org/10.1016/j.actpsy.2020.103054

Song, Y., & Hakoda, Y. (2015). An fMRI study of the functional mechanisms of Stroop/reverse-Stroop effects. *Behavioural brain research*, *290*, 187-196. https://doi.org/10.1016/j.bbr.2015.04.047

Stroop, J. R. (1992). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General*, *121*(1), 15-23. https://doi.org/10.1037/h0054651

Figure 1.

Interaction effect of task and congruence on accuracy



Note. The interaction of task and congruence on accuracy was greater for the word task than the ink task.

Image provided by Dr. Barnhardt

Table 1

Main effects and interaction effect of task and congruence on accuracy

		ANOVA				
		F ratio	*df*	η_p^2	*MSE*	*p-value*
	Task	17.30	1, 43	.29	.002	<.001
	Congruence	46.77	1, 43	.51	.002	<.001
Effect	Task x Congruence	14.28	1, 43	.25	.003	<.011

Note. *N* = 44. ANOVA = analysis of variance. Alpha level = 95%.