Emotion recognition and inhibitory control in manifest and pre-manifest Huntington's disease: evidence from a new Stroop task

Thomas Hünefeldt[1, *], Sabrina Maffi[2], Simone Migliore[2], Ferdinando Squitieri[2], Marta Olivetti Belardinelli[1, *]
1 Sapienza University of Rome, Rome, Italy
2 Huntington and Rare Diseases Unit, Fondazione IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy

Abstract

Huntington's disease (HD) is a genetic neurodegenerative disorder that affects not only the motor but also the cognitive domain. In particular, cognitive symptoms such as impaired executive skills and deficits in recognizing other individuals' mental state may emerge many years before the motor symptoms. This study was aimed at testing two cognitive hypotheses suggested by previous research with a new Stroop task created for the purpose: 1) the impairment of emotion recognition in HD is moderated by the emotions' valence, and 2) inhibitory control is impaired in HD. Forty manifest and 20 pre-manifest HD patients and their age- and gender-matched controls completed both the traditional "Stroop Color and Word Test" (SCWT) and the newly created "Stroop Emotion Recognition under Word Interference Task" (SERWIT), which consist in 120 photographs of sad, calm, or happy faces with either congruent or incongruent word interference. On the SERWIT, impaired emotion recognition in manifest HD was moderated by emotion type, with deficits being larger in recognizing sadness and calmness than in recognizing happiness, but it was not moderated by stimulus congruency. On the SCWT, six different interference scores yielded as many different patterns of group effects. Overall our results corroborate the hypothesis that impaired emotion recognition in HD is moderated by the emotions' valence, but do not provide evidence for the hypothesis that inhibitory control is impaired in HD. Further research is needed to learn more about the psychological mechanisms underlying the moderating effect of emotional valence on impaired emotion recognition in HD, and to corroborate the hypothesis that the inhibitory processes involved in Stroop tasks are not impaired in HD. Looking beyond this study, the SERWIT promises to make important contributions to disentangling the cognitive and the psychomotor aspects of neurological disorders. The research was approved by the Ethics Committee of the "Istituto Leonarda Vaccari", Rome on January 24, 2018.

Key Words: *congruent word interference; emotion recognition; emotional valence; Huntington's disease; incongruent word interference; inhibitory control; Stroop interference*

***Chinese Library Classification No.** R471; R448*

Introduction

Huntington's disease (HD) is a rare, neurodegenerative, genetic disorder with an autosomal-dominant pattern. It is caused by an expanded CAG repeat beyond 35 in the *HTT* gene on chromosome 4 (Huntington's Disease Collaborative Research Group, 1993), which leads to a mutant form of the encoded protein with a toxic function resulting in neuronal loss and abnormal brain development in childhood variants (Bordelon, 2013; Bates et al., 2015; Fusilli et al., 2018). HD affects motor, behavioral and cognitive domains, with a variable presentation of symptoms (Snowden, 2017). Motor symptoms such as balance impairment, dystonia and cho-rea are the most visible signs and they are typically used to define the onset of disease (Huntington Study Group, 1996, Marder et al., 2000). However, cognitive deficits can be present long before the onset of motor signs (Paulsen et al., 2008; Tabrizi et al., 2009, 2012, 2013; Stout et al., 2011) and con-tribute greatly to patients' loss of functional independence, with a significant impact on patients' and caregivers' quality of life (Mestre et al., 2018).

One important kind of cognitive impairment in HD con-cerns emotion recognition. In particular, there is consistent evidence that the recognition of emotions from other peo-ple's facial and non-facial expressions is impaired in both manifest and pre-manifest HD, with deficits being some-what less pronounced in pre-manifest HD (see Henley et al., 2012; Kordsachia et al., 2017 for reviews). Notably, these deficits have consistently been found to be moderated by the type of emotion that was to be recognized. In particular, a recent meta-analysis evidenced that deficits were largest with anger, disgust, and fear, somewhat smaller with sad-ness, and smallest with surprise and happiness (Bora et al., 2016). The extant findings on emotion recognition in HD thus suggest that the moderating effect of emotion type is related to differences concerning the emotions' valence, re-flecting specific impairments concerning the processing of negative emotional content in HD. In fact, this explanation is also supported by recent research on the recognition of more complex mental states in manifest and pre-manifest HD (Olivetti Belardinelli et al., 2019). However, based on the extant findings, it cannot be excluded that the moderating effect of emotion type is rather due to other stimulus-related factors such as the emotions' recognition difficulty or arous-al potential. On the one hand, in fact, emotion recognition is generally assessed by means of tests that comprise more than one negative emotion (generally anger, disgust, fear, and sadness) but only one positive emotion (happiness), so

that the moderating effect of emotion type might simply be due to differences concerning the recognition difficulty of the emotions that were to be recognized. On the other hand, the emotions for which recognition deficits in HD have been found to be largest (i.e., anger, disgust, and fear) have a particularly high arousal potential, so that the moderating effect of emotion type might as well be related to differences concerning the arousal potential of the emotions that were to be recognized. Further research is needed in order to test these different kinds of explanations.

Another important kind of cognitive impairment in HD concerns inhibitory control, shortly called inhibition, a core executive function that involves the ability to inhibit prepotent but task-irrelevant responses (Diamond, 2013; Karr et al., 2018). In fact, there seems to be substantial evidence that inhibition is impaired already in pre-manifest HD (see Rao et al., 2014 for review and additional evidence). Some studies, however, failed to find evidence of impaired inhibition in pre-manifest HD (e.g., Unmack Larsen et al., 2015) or even in manifest HD (Eddy & Rickards 2015). Part of the differences in findings might be due to the different ways in which inhibition is assessed. In fact, especially in studies focusing on other variables inhibition is assessed, sometimes exclusively, by means of the Stroop color and word test (SCWT), a specific kind of Stroop paradigm that comprises three tasks: color naming, word reading, and color naming under incongruent word interference (see below for further details). Using this test, many studies assess inhibition simply in terms of the number of correct answers in the incongruent word interference condition (e.g., Adjeroud al., 2016), whereas others assess it in terms of a composite interference score that takes into account a non-interference condition as baseline condition (e.g., Adjeroud et al., 2017) or in terms of an error rather than an accuracy score (Eddy and Rickards, 2015). Only occasionally some other kind of Stroop test is used that takes into account a congruent interference condition as baseline condition (Unschuld et al., 2012). However, performance on Stroop tasks is determined by different factors (Kane and Engle, 2003), and this is all the more so the case when the baseline condition is not a congruent interference condition but a non-interference condition, let alone when no baseline is taken into account at all. Accordingly, given that the different factors involved in performance on Stroop tasks might be differently impaired in HD, the evidence for impaired inhibition in HD might depend both on the Stroop task and on the scoring method that are used.

Aims
The aim of this study was to test predictions concerning emotion recognition and response inhibition in manifest and pre-manifest HD with a new Stroop task. In particular, we wanted to test the following two predictions suggested by previous findings: 1) the impairment of emotion recognition in HD is moderated by the emotions' valence, and 2) inhibitory control is impaired in HD.

Participants and Methods
Participants
The study sample was the same used in a previous research

(Olivetti Belardinelli et al., 2019) and consisted of 140 participants belonging to four groups: 40 participants with manifest HD (HD), 40 age-matched healthy control subjects (HC), 20 participants with pre-manifest HD (preHD), and 40 age-matched healthy control subjects (preHC). HD and preHD were recruited at the LIRH Foundation sites in Northern, Middle and Southern Italy and at Fondazione IRCCS Casa Sollievo della Sofferenza Research Hospital, CSS-Mendel Institute of Human Genetics in Rome. HC and preHC were selected from hospital employees and their relatives. None of them was biologically or personally related to HD patients.

General inclusion criteria were right-handedness, normal color vision, and normal or corrected to normal visual acuity. General exclusion criteria were medical conditions that might influence cognition (e.g., a history of developmental disorder, psychotic disorder, or substance or alcohol dependence) and incomplete test performance. Specific inclusion criteria for HD and preHD were a positive genetic test for HD with a CAG expansion ≥ 40. Following an established diagnostic standard (Paulsen et al., 2001), the distinction between HD and preHD was defined on the basis of the total motor score and diagnostic confidence level, with total motor scores > 10 and diagnostic confidence level = 4 for HD, and total motor scores ≤ 10 and diagnostic confidence level < 4 for preHD.

Mann-Whitney U tests indicated that the age differences between HD and HC ($U = 704.5$, $Z = –0.91$, $P = 0.36$) and between preHD and preHC ($U = 331$, $Z = –1.07$, $P = 0.28$) were not statistically significant. The total functioning capacity, assessing the capacity to perform a range of activities of basic daily living (i.e., working, chores, managing finances, eating, dressing and bathing), was significantly lower in HD than in preHD ($P < 0.001$). Basic demographic and clinical information on the four groups of participants is summarized in **Table 1**.

Measures
Stroop emotion recognition under word interference task
The Stroop emotion recognition under word interference task (SERWIT) is an emotion recognition task that we constructed for the scope of this research. It consists in 120 photographs of happy, calm ("neutral") or sad male or female

Table 1 Basic demographic and clinical characteristics of participants with manifest HD (HD), pre-manifest HD (preHD), and their respective controls (HC and preHC)

	HD	HC	preHD	preHC
Sex (M/F, *n*)	40 (20/20)	40 (20/20)	20 (10/10)	40 (20/20)
Age (years)				
Mean±SD	45.3±10.1	47.2±7.2	34.9±8.9	31.7±4.7
Range	27–69	40–67	22–48	23–39
TMS				
Mean±SD	32.9±12.4	–	5.3±2.3	–
Range	11–57	–	1–10	–
TFC				
Mean±SD	9.8±1.8	–	12.9±0.7	–
Range	7–13	–	10–13	–

F: Female; M: male; TFC: total functioning capacity; TMS: total motor score.

faces, taken from the "Radboud Faces Database" (Langner et al., 2010), with the emotion words "HAPPY", "CALM" or "SAD" written in capital letters on the nose region of each face (**Figure 1**). The stimuli are balanced by expressor gender (male, female), emotion type (sadness, calmness, happiness), and stimulus congruency (congruent, incongruent) (**Table 2**). In congruent stimuli, the emotion word matches the emotion expressed in the photograph (e.g., the word "HAPPY" on a happy face), whereas in incongruent stimuli it does not (e.g., the word "HAPPY" on a sad face). Participants are requested to recognize the emotion expressed on the face. Performance is assessed both in terms of the number of correctly recognized items and in terms of the mean response time.

Stroop color and word test
The SCWT is part of the Unified Huntington's Disease Rating Scale (Huntington Study Group, 1996) and consists in three tasks. In the color naming task (*C*), participants are requested to name the color of randomly presented tokens of the three colors blue, red, and green. In the word reading
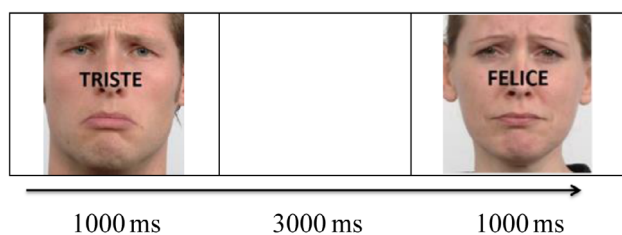
Figure 1 Example of the SERWIT stimuli and procedure.
The Italian words "TRISTE" and "FELICE" mean "SAD" and "HAPPY", respectively (as faces in Figure 1 were obtained from a public database (http://www.socsci.ru.nl:8180/RaFD2/RaFD), there is no need to de-identify it).

Table 2 Distribution of the SERWIT stimuli by emotion type (sad, calm, and happy), stimulus congruency (congruent and incongruent), and expressor gender (male and female)

Stimulus		Stimulus congruency	Expressor gender	Number of stimuli
Facial emotion	Emotion word			
Sad	Sad	Congruent	Male	10
	Sad		Female	10
	Calm	Incongruent	Male	5
	Calm		Female	5
	Happy	Incongruent	Male	5
	Happy		Female	5
Calm	Sad	Incongruent	Male	5
	Sad		Female	5
	Calm	Congruent	Male	10
	Calm		Female	10
	Happy	Incongruent	Male	5
	Happy		Female	5
Happy	Sad	Incongruent	Male	5
	Sad		Female	5
	Calm	Incongruent	Male	5
	Calm		Female	5
	Happy	Congruent	Male	10
	Happy		Female	10

task (*W*), participants are requested to read randomly presented tokens of the three color words "BLUE", "RED", and "GREEN" printed in black ink. Finally, in the color-word-interference task (*CW*), participants are requested to name the color of randomly presented tokens of the three color words "BLUE", "RED", or "GREEN" printed in an inconsistent color ink (e.g., the word "BLUE" is printed either in red or in green). Based on the answers given in 45 seconds on each of these three tasks, three different scores are obtained for each task: the number of correctly named items, the number of errors, and the number of self-corrected errors. For the scope of this study, we considered the following six interference scores:

- The interference accuracy score (*CW*) is the most frequently used interference score in research on HD (e.g., Harrington et al., 2012; Rao et al., 2014). It consists in the number of correctly named items in the color-word-interference task.
- The interference error score (CW_{err}) has occasionally been used in research on HD (e.g. Snowden et al., 2002). It consists in the number of errors (including self-corrected errors) in the color-word-interference task.
- The interference difference accuracy score (*CW-C*) is a composite interference score that has occasionally been used in studies on HD (e.g., Adjeroud et al., 2017). It consists in the difference between the number of correctly named items in the color-word-interference task (*CW*) and the number of correctly named items in the color naming task (*C*):
$$CW\text{-}C = CW - C$$
- The interference difference error score ($CW\text{-}C_{err}$) is a composite interference score that has occasionally been used in studies on HD (e.g., Eddy and Rickards, 2015). It consists in the difference between the number of errors (including self-corrected errors) in the color-word-interference task (CW_{err}) and the number of errors (including self-corrected errors) in the color naming task (C_{err}):
$$CW\text{-}C_{err} = CW_{err} - C_{err}$$
- The interference ratio score (*CW/C*) is a composite interference score that has repeatedly been used in normative studies on the SCWT, but to the best of our knowledge it has not been used so far in research on HD. It consists in the ratio between the number of correctly named colors in the color-word-interference task (*CW*) and the number of correctly named colors in the color naming task (*C*):
$$CW/C = CW \div C$$
- Golden's (1978) "interference index" (IG) is by far the most widely used score in normative studies on the SCWT (Scarpina and Tagini 2017), but to the best of our knowledge it has not been used so far in research on HD. It consists in the difference between *CW* and a theoretically predicted *CW* score (*PCW*):
$$IG = CW - PCW, \text{ with } PCW = (C \times W) \div (C + W)$$

Procedures
After giving their informed consent, participants received first the SCWT, then a mindreading task (Olivetti Belardinelli et al., 2019), and finally the SERWIT. Following the standard procedure, the SCWT was administered as a pa-

per-pencil test. The SERWIT was administered on a large laptop monitor (15.6 inches). The SERWIT stimuli were randomly presented for 1000 ms, followed by a blank slide presented for 3000 ms (**Figure 1**). Before starting the SERWIT, all participants had a test trial with 10 stimuli.

The research was completed in accordance with the *Declaration of Helsinki* and with a procedure approved by the Ethics Committee of the "Istituto Leonarda Vaccari", Rome on January 24, 2018. All participants were tested at the Sapienza University of Rome between February 2018 and April 2018 and signed a written informed consent.

Design and analyses
We performed three analysis of variance (ANOVA) with group (HD, HC, preHD, preHC) as between-subject variable: a multivariate ANOVA on the eight SCWT scores (*C, W, CW CW*$_{err}$, *CW-C, CW-C*$_{err}$, *CW/C*, IG), and two repeated-measures ANOVA with stimulus congruency (congruent, incongruent) and stimulus emotion type (sadness, calmness, and happiness) as within-subject variables on the SERWIT scores concerning emotion recognition accuracy and response time, respectively. All statistical analyses were performed using IBM SPSS Statistics 19.0 (IBM, Armonk, NY, USA).

Results
Effects on the SERWIT
Effects on recognition accuracy
The repeated-measures ANOVA with group (HD, HC, preHD, and preHC) as between-subject variable and stimulus congruency (incongruent, congruent) and stimulus emotion type (sadness, calmness, and happiness) as within-subject variables revealed significant main effects of group, stimulus congruency, and stimulus emotion type, on emotion recognition accuracy. Furthermore, there were significant two-way interactions between group and stimulus emotion type, and between stimulus congruency and stimulus emotion type, whereas the two-way interaction between group and stimulus congruency, and the three-way interaction between group, stimulus congruency, and stimulus emotion type, were far from being statistically significant. Details are reported in **Table 3**.

Table 3 Results of the repeated-measures analysis of variance with group (G) as between-subject variable and stimulus congruency (SC) and stimulus emotion type (SET) as within-subject variables on the SERWIT emotion recognition accuracy

Factor	df	Error	F	η_p^2
G	3	136	34.8***	0.43
SC	1	136	36.7***	0.21
SET	2	272	69.2***	0.34
G•SC	3	136	0.4	0.01
G•SET	6	272	5.9***	0.12
SC•SET	2	272	11.4***	0.08
G•SC•SET	6	272	0.6	0.01

***P < 0.001. SERWIT: Stroop emotion recognition under word interference task.

As to the significant main effect of group, emotion recognition accuracy was overall significantly lower in HD (mean = 84.0, standard error = 2.1) than in HC (mean = 110.0, standard error = 2.1, *P* < 0.001), whereas it did not differ significantly between preHD (mean = 102.9, standard error = 2.9) and preHC (mean = 109.7, standard error = 2.1, *P* = 0.4). Furthermore, recognition accuracy was overall significantly lower in HD than in preHD (*P* < 0.001), whereas it did not differ significantly between HC and preHC (*P* = 1).

As to the significant main effect of stimulus congruency, emotion recognition accuracy was overall significantly higher for emotions displayed together with congruent emotion terms (mean = 104.8, standard error = 1.3) than for emotions displayed together with incongruent emotion terms (mean = 98.5, standard error = 1.2).
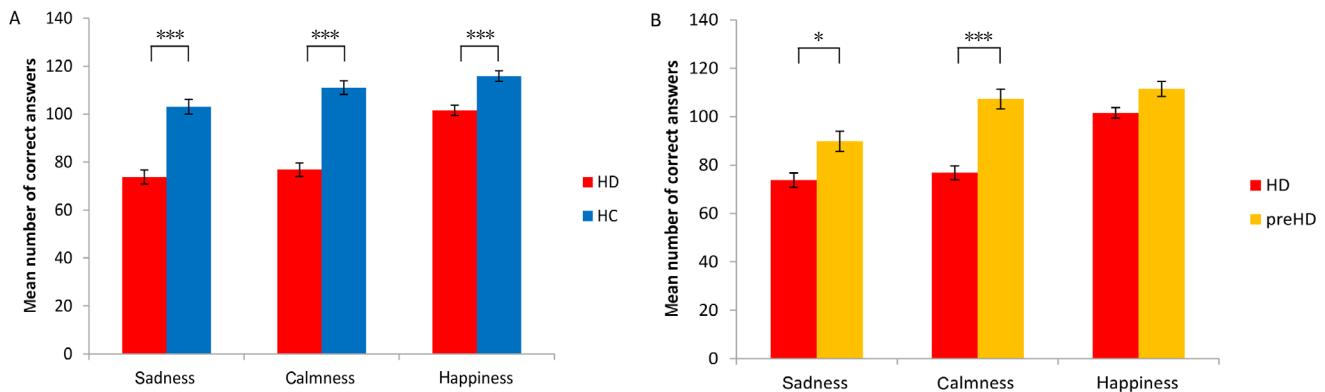
As to the significant main effect of stimulus emotion type, emotion recognition accuracy was overall significantly higher for happiness (mean = 111.4, standard error = 1.2) than for sadness (mean = 91.8, standard error = 1.7, *P* < 0.001) and calmness (mean = 101.6, standard error = 1.6, *P* < 0.001), and it was also significantly higher for calmness than for sadness (*P* < 0.001).

As to the significant interaction between stimulus congruency and stimulus emotion type, follow-up analyses revealed effects of stimulus congruency only concerning sadness, $F_{(1, 139)}$ = 31.3, *P* < 0.001, η_p^2 = 0.18, and calmness, $F_{(1, 139)}$ = 31.3, *P* < 0.001, η_p^2 = 0.18, but not concerning happiness, $F_{(1, 139)}$ = 1.2, *P* = 0.3, η_p^2 = 0.01. In particular, emotion recognition accuracy was significantly higher for congruent than for incongruent stimuli in case of sadness (congruent: mean = 96.4, standard error = 2.2, incongruent: mean = 87.8, standard error = 1.8) and in case of calmness (congruent: mean = 105.1, standard error = 2.0, incongruent: mean = 96.5, standard error = 2.2), but not in case of happiness (congruent: mean = 112.0, standard error = 1.4, incongruent: mean = 110.8, standard error = 1.4).

Finally, as to the significant interaction between group and stimulus emotion type, follow-up analyses confronting each of the four groups of participants with each other revealed that the interaction between group and stimulus emotion type was significant both when confronting HD and HC, $F_{(2, 156)}$ = 9.2, *P* < 0.001, η_p^2 = 0.11, and when confronting HD and preHD, $F_{(2, 116)}$ = 4.8, *P* < 0.05, η_p^2 = 0.08, but not when confronting preHD and preHC, $F_{(2, 116)}$ = 1.6, *P* = 0.2, η_p^2 = 0.03, nor when confronting preHC and HC, $F_{(2, 156)}$ = 0.8, *P* = 0.5, η_p^2 = 0.01. In particular, HD recognized all three types of emotions significantly less accurately than HC, but the difference between HD and HC was larger in recognizing sadness and calmness than in recognizing happiness (**Figure 2A**). Similarly, HD recognized all three types of emotions less accurately than preHD, but the difference between HD and preHD was statistically significant only in recognizing sadness and calmness, but not in recognizing happiness (**Figure 2B**).

Effects on response time
The repeated-measures ANOVA with group (HD, HC, preHD, and preHC) as between-subject variable and stimulus congruency (incongruent and congruent) and stimulus emo-

Figure 2 Emotion recognition accuracy for sadness, calmness, and happiness on the SERWIT in participants with manifest HD (HD) compared with healthy controls (HC; A) and participants with pre-manifest HD (preHD; B).
$*P < 0.05$, $***P < 0.001$.

tion type (sadness, calmness, and happiness) as within-subject variables revealed significant main effects of group, stimulus congruency, and stimulus emotion type, whereas none of the interactions between these variables was statistically significant. Details are reported in **Table 4**.

As to the significant main effect of group, response times were overall significantly longer in HD (mean = 1.68, standard error = 0.07) than in HC (mean = 0.96, standard error = 0.07, $P < 0.001$), whereas they did not differ significantly between preHD (mean = 1.01, standard error = 0.10) and preHC (mean = 0.85, standard error = 0.07, $P = 1$). Furthermore, response times were overall significantly longer in HD than in preHD ($P < 0.001$), whereas it did not differ significantly between HC and preHC ($P = 1$).

As to the significant main effect of stimulus congruency, response times were significantly shorter for emotions displayed together with congruent emotion terms (mean = 1.11, standard error = 0.04) than for emotions displayed together with incongruent emotion terms (mean = 1.14, standard error = 0.04).

Finally, as to the significant main effect of stimulus emotion type, response times were significantly shorter for happiness (mean = 1.03, standard error = 0.04) than for sadness (mean = 1.19, standard error = 0.04, $P < 0.001$) and calmness (mean = 1.15, standard error = 0.04, $P < 0.001$), whereas they did not differ significantly between sadness and calmness ($P = 0.08$).

Table 4 Results of the repeated-measures analysis of variance with group (G) as between-subject variable and stimulus congruency (SC) and stimulus emotion type (SET) as within-subject variables on the SERWIT response time

Factor	df	Error	F	η_p^2
G	3	136	29.8***	0.40
SC	1	136	6.0*	0.04
SET	2	272	48.6***	0.26
G•SC	3	136	0.2	0.00
G•SET	6	272	1.9	0.04
SC•SET	2	272	1.7	0.00
G•SC•SET	6	272	0.7	0.01

$*P < 0.05$, $***P < 0.001$. SERWIT: Stroop emotion recognition under word interference task.

Effects on the SCWT

The one-way multivariate ANOVA with group (HD, HC, preHD, and preHC) as between-subject variable revealed significant main effects on all SCWT scores except $CW\text{-}C_{err}$ (**Table 5**). Details on group differences are reported in **Table 6**.

Table 5 Results of the multivariate analysis of variance with group as between-subject variable on the eight SCWT scores (C, W, CW, CW_{err}, $CW\text{-}C$, $CW\text{-}C_{err}$, CW/C, and IG)

Dependent variable	df	Error	F	η_p^2
C	3	136	83.3***	0.65
W	3	136	98.8***	0.69
CW	3	136	87.9***	0.66
CW_{err}	3	136	5.7**	0.11
$CW\text{-}C$	3	136	9.7***	0.18
$CW\text{-}C_{err}$	3	136	0.8	0.02
CW/C	3	136	6.2**	0.12
IG	3	136	8.8***	0.16

$*P < 0.05$, $***P < 0.001$. C: Color naming task; CW: color-word-interference task; $CW\text{-}C$: interference difference accuracy score; CW/C: interference ratio score; $CW\text{-}C_{err}$: interference difference error score; CW_{err}: interference error score; IG: interference index; SCWT: Stroop color and word test; W: word reading task.

Table 6 Means and standard errors for the eight SCWT scores (C, W, CW, CW_{err}, $CW\text{-}C$, $CW\text{-}C_{err}$, CW/C, and IG) of participants with manifest HD (HD), pre-manifest HD (preHD), and their respective controls (HC and preHC)

	HD	HC	preHD	preHC
C	45.4 (1.9)***###	76.1 (1.9)	71.2 (2.7)***###	86.1 (1.9)
W	63.4 (2.4)***###	104.8 (2.4)	96.9 (3.3)***###	118.4 (2.4)
CW	24.6 (1.4)***###	45.9 (1.4)	40.2 (1.9)***###	55.3 (1.4)
CW_{err}	2.20 (0.26)	1.30 (0.26)	1.55 (0.37)	0.68 (0.26)
$CW\text{-}C$	−20.9 (1.5)***###	−30.2 (1.5)	−31.0 (2.2)##	−30.7 (1.5)
$CW\text{-}C_{err}$	0.98 (0.26)	0.58 (0.26)	0.50 (0.34)	0.43 (0.26)
CW/C	0.54 (0.02)*	0.61 (0.02)	0.57 (0.04)	0.65 (0.02)
IG	−1.75 (1.08)	2.00 (1.08)	−0.79 (1.53)**	5.67 (1.08)

Pairwise comparisons between HD and HC, and between preHD and preHC: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$; pairwise comparisons between HD and preHD: $\#\#P < 0.01$, $\#\#\#P < 0.001$. C: color naming task; CW: color-word-interference task; $CW\text{-}C$: interference difference accuracy score; CW/C: interference ratio score; $CW\text{-}C_{err}$: interference difference error score; CW_{err}: interference error score; IG: interference index; SCWT: Stroop color and word test; W: word reading task.

As to the significant main effects on *C*, *W*, and *CW*, on all three tasks HD performed significantly worse than both HC ($P < 0.001$, on all three tasks) and preHD ($P < 0.001$, on all three tasks), and preHD performed significantly worse than preHC ($P < 0.001$, on all three tasks), with preHC performing significantly better than HC ($P < 0.01$ for *C*, $P < 0.001$ for *W*, and $P < 0.001$ for *CW*).

As to the significant main effect on CW_{err}, the only significant difference concerned HD and preHC ($P < 0.001$). The differences between HD and HC ($P = 0.1$), HD and preHD ($P = 0.9$), preHD and preHC ($P = 0.3$), and preHC and HC ($P = 0.6$), were all not significant.

As to the significant main effect on *CW-C*, the difference between *CW* and *C* was significantly smaller for HD than for HC ($P < 0.001$) and preHD ($P < 0.01$), but it did not differ significantly neither between preHD and preHC ($P = 1$) nor between preHC and HC ($P = 1$).

As to the significant main effect on *CW/C*, the ratio between *CW* and *C* tended to be significantly smaller in HD than in HC ($P = 0.05$), but it did not differ significantly neither between preHD and preHC ($P = 1$) nor between preHC and HC ($P = 0.08$).

Finally, as to the significant main effect on IG, the interference index was significantly smaller for preHD than for preHC ($P < 0.01$), but did not differ significantly neither between HD and HC ($P = 0.09$) nor between HD and preHD ($P = 1$) nor between HC and preHC ($P = 0.1$).

Discussion

The aim of this study was to test predictions concerning emotion recognition and inhibitory control in manifest and pre-manifest HD with a new Stroop task. The SERWIT is an easy-to-administer task that resembles but also differs from the SCWT in important respects. It resembles the SCWT in being a Stroop task involving a word interference condition, but unlike the SCWT it does not assess color naming in an incongruent word interference condition but emotion recognition in both a congruent and an incongruent word interference condition. Unlike the SCWT, it therefore allows assessing inhibition independently from other factors involved in Stroop performance. Furthermore, unlike other mental state recognition tasks, it allows simultaneously controlling for different stimulus-related variables such as expressor gender, emotional valence, and arousal potential. Given these general task characteristics, we will now discuss our results with respect to the two predictions motivated in our Introduction.

Is the impairment of emotion recognition in HD moderated by the emotions' valence?

The fact that we found significant interactions between group and emotion type, with HD being less impaired in recognizing happiness than in recognizing sadness and calmness, suggests that HD's impairment in emotion recognition is moderated by the type of emotions. The most straightforward explanation of this moderating effect of emotion type is in terms of emotional valence, with HD being less impaired in recognizing positive emotions such as happiness than rather neutral or negative emotions such as calmness and sadness. In fact, alternative explanations in terms of arousal potential or recognition difficulty are less parsimonious and overall rather unlikely.

On the one hand, in fact, explanations in terms of arousal potential would have to suppose that happiness is either more or less arousing than both calmness and sadness. However, this seems much unlikely, given that emotionally positive and negative stimuli are known to be more arousing than emotionally neutral ones. On the other hand, explanations in terms of recognition difficulty would have to suppose that happiness is easier to recognize than both calmness and sadness. However, sadness, calmness, and happiness are all rather simple emotions that are similarly easy to recognize, and they were the only emotions that were to be recognized. Unlike in other studies, the difficulty of recognizing sadness *vs.* happiness was thus not affected by the presence of other negative emotions. Furthermore, differences in recognition difficulty might themselves well be due to differences in their emotional valence. Therefore, the moderating effect of emotion type is overall more likely due to differences in emotional valence than to differences in recognition difficulty.

Supposed, then, that the impairment of emotion recognition in HD is moderated by the emotions' valence, there are basically two different but complementary kinds of hypotheses concerning the psychological mechanism underlying this moderating effect. On the one hand, one might hypothesize that HD patients tend more strongly to avoid processing emotionally negative stimuli, for example for fear of being confronted with their own distressful condition (Olivetti Belardinelli et al., 2019). On the other hand, one might hypothesize that HD tend more strongly to process emotionally positive stimuli, for example in order to get some relief and distraction from their condition. Given that in our study HD's impairment was particularly strong not only with sadness but also with calmness, the former hypothesis seems less likely than the latter. However, it might well be that the particular impairment pattern is due to overlapping effects of emotional valence, arousal potential, and recognition difficulty. Therefore, our study does not allow any definite conclusion as to which of the two hypotheses is more plausible.

Is inhibitory control impaired in HD?

The fact that on the SERWIT we found no significant interaction between group and stimulus congruency concerning the accuracy and speed of emotion recognition suggests that inhibition is impaired neither in pre-manifest nor in manifest HD. This conclusion seems to contradict not only the results of most of the previous studies with the SCWT, but also some of the results obtained in this study. In fact, on four of the six different SCWT interference scores that we considered in this study, we found significant group differences between HD and HC, preHD and preHC, or preHD and HD. Notably, however, these four SCWT interference scores yielded four substantially different patterns of group effects. First, as regards *CW*, all differences between groups were significant. Second, as regards *CW-C*, HD differed significantly from both HC and preHD, but preHD did not differ significantly from preHC. Third, as regards *CW/C*, HD

tended to differ significantly from HC but not from preHD, and preHD did not differ significantly from preHC. Finally, as regards the IG, the only significant difference was between preHD and preHC. Upon closer consideration, these four heterogeneous results provide only very little if any evidence of impaired inhibition in HD:

First, as regards *CW*, lower *CW* scores do not necessarily imply impaired inhibition. In fact, differences in the number of colors correctly named in the incongruent interference condition might easily be explained in terms of other factors such as overall processing speed. Suppose, for example, that overall processing speed but neither inhibition nor any other variable is impaired in HD. Then the difference between the *CW* scores of HD and HC is the following function of the difference between the processing speeds of HD and HC: $CW_{HD} - CW_{HC} = IC \times (V_{HD} - V_{HC}) \times T$, where *IC* (with $0 \leq IC \leq 1$) is an interference coefficient, *V* the processing speed in the non-interference condition, and *T* the test duration. Thus, the difference between the *CW* scores of HD and HC can theoretically exceed any limit, provided that the difference between processing speeds is sufficiently large and the test duration is sufficiently long.

Second, as regards the interference difference accuracy score (*CW-C*), the absolute value of the difference between the scores in the incongruent interference condition and the scores in the non-interference condition was significantly smaller for HD than for HC. This remarkable finding, which is in line with the results of other studies that take into account a non-interference condition as baseline condition (e.g., Adjeroud et al., 2017), suggests either that inhibition was stronger in HD than in HC or that other factors are responsible for the observed effect. For example, it might easily be explained in terms differences in overall processing speed. In fact, the difference between the scores in the incongruent interference condition and the scores in the non-interference condition may be described as the following function of processing speed: $CW-C = (IC-1) \times V \times T$, where *IC* (with $0 \leq IC \leq 1$) is an interference coefficient, *V* the processing speed in the non-interference condition, and *T* the test duration. Accordingly, the absolute value of *CW-C* scores may be expected to be the smaller the smaller is the processing speed, and the difference between the *CW-C* scores of HD and HC can theoretically exceed any limit, provided that the difference between their processing speeds is sufficiently large and the test duration is sufficiently long.

Third, as regards *CW/C*, the fact that *CW/C* tended to be significantly lower in HD than in HC provides only very weak evidence of impaired inhibition in HD. In fact, even if this only marginally significant effect should be reliable, it need not be related to differences in inhibition, because it does not concern the relation between an incongruent and a congruent interference condition but the relation between an incongruent interference and a non-interference condition. Though it cannot be accounted for in terms of differences concerning processing speed, it might therefore as well be related to other factors involved in Stroop performance.

Finally, as regards the IG, the pattern of results concerning HD and preHD implies that neither in HD nor in preHD the IG score was significantly below zero. As IG consists in the difference between the observed *CW* score and a theoretically predicted *CW* score, this implies that the *CW* score was neither in HD nor in preHD significantly lower than the theoretically predicted *CW* score. Accordingly, our results concerning IG do not provide evidence of impaired inhibition.

To sum up, most of the significant group effects found with the SCWT can be explained more parsimoniously in terms of factors other than inhibition, whereas most of the group effects that cannot be easily explained in terms of other factors were not statistically significant or otherwise don't provide evidence of impaired inhibition in HD. Given that our findings with the SERWIT also did not provide evidence of impaired inhibition in HD, our results therefore overall provide converging evidence that the inhibitory processes involved in performance on Stroop tasks are not impaired in HD. Notably, our results do in no way exclude that the inhibitory processes involved in performance on other kinds of tasks may be impaired in HD. They therefore do not contradict the results of studies using other kinds of test to assess inhibitory control (e.g., Henderson et al., 2011). They rather suggest that inhibitory control is a rather complex construct and that not all of its aspects are impaired in HD.

Study limitations

A general limitation of our study consists in the fact that participants were not requested to suspend their pharmacological treatments before testing. Accordingly, the test performance especially of the two patient groups may have been somewhat different from what it would have been without medication. Furthermore, there are mainly two more specific limitations. First, as regards the moderating effect of emotional valence on emotion recognition, we did not assess whether the stimuli representing happiness, calmness, and sadness differed in their arousal potential and recognition difficulty. Our findings therefore do not exclude the possibility that the effect of emotional valence is caused by differences in arousal potential or recognition difficulty, even though we have provided arguments as to why this possibility is rather unlikely. Second, we did not use another test of inhibitory control that is not based on a Stroop paradigm. Our findings therefore do not allow conclusions concerning the inhibitory processes involved in performance on other kinds of tasks.

Conclusions

This study was aimed at testing two predictions concerning emotion recognition and inhibitory control in manifest and pre-manifest HD by means of a new Stroop task called SERWIT. On the one hand, our results corroborate the hypothesis that the impairment of emotion recognition in HD is moderated by the emotions' valence. On the other hand, our results cast don't provide evidence for the hypothesis that inhibitory control is impaired in HD. However, further research is needed to learn more about the psychological mechanisms underlying the moderating effect of emotional valence on impaired emotion recognition in HD, and to corroborate the hypothesis that the inhibitory processes involved in Stroop performance are not impaired in HD. Looking beyond this study, the SERWIT promises to make

important contributions to disentangling the cognitive and the psychomotor aspects of neurological disorders.

References

Adjeroud N, Besnard J, El Massioui N, Verny C, Prudean A, Scherer C, Gohier B, Bonneau D, Allain P (2016) Theory of mind and empathy in preclinical and clinical Huntington's disease. Soc Cogn Affect Neurosci 11:89-99.

Adjeroud N, Besnard J, Verny C, Prundean A, Scherer C, Gohier B, Bonneau D, Massioui NE, Allain P (2017) Dissociation between decision-making under risk and decision-making under ambiguity in premanifest and manifest Huntington's disease. Neuropsychologia 103:87-95.

Bates GP, Dorsey R, Gusella JF, Hayden MR, Kay C, Leavitt BR, Nance M, Ross CA, Scahill RI, Wetzel R, Wild EJ, Tabrizi SJ (2015) Huntington disease. Nat Rev Dis Primers 1:15005.

Bora E, Velakoulis D, Walterfang M (2016) Social cognition in Huntington's disease: a meta-analysis. Behav Brain Res 297:131-140.

Bordelon YM (2013) Clinical neurogenetics: huntington disease. Neurol Clin 31:1085-1094.

Diamond A (2013) Executive functions. Annu Rev Psychol 64:135-168.

Eddy CM, Rickards HE (2015) Cognitive deficits predict poorer functional capacity in Huntington's disease: but what is being measured? Neuropsychology 29:268-273.

Fusilli C, Migliore S, Mazza T, Consoli F, De Luca A, Barbagallo G, Ciammola A, Gatto EM, Cesarini M, Etcheverry JL, Parisi V, Al-Oraimi M, Al-Harrasi S, Al-Salmi Q, Marano M, Vonsattel JG, Sabatini U, Landwehrmeyer GB, Squitieri F (2018) Biological and clinical manifestations of juvenile Huntington's disease: a retrospective analysis. Lancet Neurol 17:986-993.

Golden CJ (1978) Stroop color and word test: a manual for clinical and experimental uses. Chicago, IL: Stoelting Co.

Henderson T, Georgiou-Karistianis N, White O, Millist L, Williams DR, Churchyard A, Fielding J (2011) Inhibitory control during smooth pursuit in Parkinson's disease and Huntington's disease. Mov Disord 26:1893-1899.

Henley SM, Novak MJ, Frost C, King J, Tabrizi SJ, Warren JD (2012) Emotion recognition in Huntington's disease: a systematic review. Neurosci Biobehav Rev 36:237-253.

Huntington's Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 72:971-983.

Huntington Study Group (1996) Unified Huntington's disease rating scale: reliability and consistency. Mov Disord 11:136-142.

Karr JE, Areshenkoff CN, Rast P, Hofer SM, Iverson GL, Garcia-Barrera MA (2018) The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. Psychol Bull 144:1147-1185.

Kordsachia CC, Labuschagne I, Stout JC (2017) Beyond emotion recognition deficits: A theory guided analysis of emotion processing in Huntington's disease. Neurosci Biobehav Rev 73:276-292.

Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, van Knippenberg A (2010) Presentation and validation of the Radboud Faces Database. Cogn Emot 24:1377-1388.

Kane MJ, Engle RW (2003) Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. J Exp Psychol Gen 132:47-70.

Marder K, Zhao H, Myers RH, Cudkowicz M, Kayson E, Kieburtz K, Orme C, Paulsen J, Penney JB Jr, Siemers E, Shoulson I (2000) Rate of functional decline in Huntington's disease. Huntington Study Group. Neurology 54:452-458.

Mestre TA, Carlozzi NE, Ho AK, Burgunder JM, Walker F, Davis AM, Busse M, Quinn L, Rodrigues FB, Sampaio C, Goetz CG, Cubo E, Martinez-Martin P, Stebbins GT; Members of the MDS Committee on Rating Scales Development (2018) Quality of life in Huntington's disease: critique and recommendations for measures assessing patient health-related quality of life and caregiver quality of life. Mov Disord 33:742-749.

Olivetti Belardinelli M, Huenefeldt T, Maffi S, Squitieri F, Migliore S (2019) Effects of stimulus-related variables on mental states recognition in Huntington's disease. Int J Neurosci 129:563-572.

Paulsen JS, Zhao H, Stout JC, Brinkman RR, Guttman M, Ross CA, Como P, Manning C, Hayden MR, Shoulson I, Huntington Study Group (2001) Clinical markers of early disease in persons near onset of Huntington's disease. Neurology 57:658-662.

Paulsen JS, Langbehn DR, Stout JC, Aylward E, Ross CA, Nance M, Guttman M, Johnson S, MacDonald M, Beglinger LJ, Duff K, Kayson E, Biglan K, Shoulson I, Oakes D, Hayden M; Predict-HD Investigators and Coordinators of the Huntington Study Group (2008) Detection of Huntington's disease decades before diagnosis: the Predict-HD study. J Neurol Neurosurg Psychiatry 79:874-880.

Rao JA, Harrington DL, Durgerian S, Reece C, Mourany L, Koenig K, Lowe MJ, Magnotta VA, Long JD, Johnson HJ, Paulsen JS, Rao SM (2014) Disruption of response inhibition circuits in prodromal Huntington disease. Cortex 58:72-85.

Scarpina F, Tagini S (2017) The Stroop Color and Word Test. Front Psychol 8:557.

Snowden JS (2017) The neuropsychology of Huntington's disease. Arch Clin Neuropsychol 32:876-887.

Snowden JS, Craufurd D, Thompson J, Neary D (2002) Psychomotor, executive and memory function in preclinical Huntington's disease. J Clin Exp Neuropsychol 24:133-145.

Stout JC, Paulsen JS, Queller S, Solomon AC, Whitlock KB, Campbell JC, Aylward EH (2011) Neurocognitive signs in prodromal Huntington disease. Neuropsychology 25:1-14.

Tabrizi SJ, Langbehn DR, Leavitt BR, Roos RA, Durr A, Craufurd D, Kennard C, Hicks SL, Fox NC, Scahill RI, Borowsky B, Tobin AJ, Rosas HD, Johnson H, Reilmann R, Landwehrmeyer B, Stout JC; TRACK-HD investigators (2009) Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: cross-sectional analysis of baseline data. Lancet Neurol 8:791-801.

Tabrizi SJ, Reilmann R, Roos RA, Durr A, Leavitt B, Owen G, Jones R, Johnson H, Craufurd D, Hicks SL, Kennard C, Landwehrmeyer B, Stout JC, Borowsky B, Scahill RI, Frost C, Langbehn DR; TRACK-HD investigators (2012) Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: analysis of 24 month observational data. Lancet Neurol 11:42-53.

Tabrizi SJ, Scahill RI, Owen G, Durr A, Leavitt BR, Roos RA, Borowsky B, Landwehrmeyer B, Frost C, Johnson H, Craufurd D, Reilmann R, Stout JC, Langbehn DR; TRACK-HD Investigators (2013) Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: analysis of 36-month observational data. Lancet Neurol 12:637-649.

Unmack Larsen I, Vinther-Jensen T, Gade A, Nielsen JE, Vogel A (2015) Assessing impairment of executive function and psychomotor speed in premanifest and manifest Huntington's disease gene-expansion carriers. J Int Neuropsychol Soc 21:193-202.

Unschuld PG, Joel SE, Pekar JJ, Reading SA, Oishi K, McEntee J, Shanahan M, Bakker A, Margolis RL, Bassett SS, Rosenblatt A, Mori S, van Zijl PC, Ross CA, Redgrave GW (2012) Depressive symptoms in prodromal Huntington's Disease correlate with Stroop-interference related functional connectivity in the ventromedial prefrontal cortex. Psychiatry Res 203(2-3):166-174.