EMOTION RECOGNITION IS IMPAIRED ACROSS MODALITIES IN MANIFEST HUNTINGTON'S DISEASE

Susie M.D. Henley

D.Clin.Psy. Thesis (Volume 1), 2011

University College London

OVERVIEW

This thesis aims to address some of the unresolved issues around emotion recognition in Huntington's disease (HD). HD is an autosomal dominant neurodegenerative disease, with symptoms usually occurring in the middle decades of life and death following approximately 20 years later. Symptoms include involuntary movements (chorea), psychiatric disturbance (e.g. irritability, apathy), and cognitive difficulties, including difficulties recognising emotions.

Part 1 of the thesis systematically reviews existing literature looking at emotion recognition in HD. Since the late 90s this has been a topic of some interest, and in particular debate has centred around whether specific emotions are disproportionately affected, and the implications of this for the neural substrates of emotion recognition given the underlying neuropathology of the disease. Therefore Part 1 aimed to assess systematically the findings to date, summarising outcomes and relevant methodological issues, and making recommendations for future research.

Part 2 follows on from some of the recommendations made in Part 1. Firstly, it examines emotion recognition in more than one modality in one HD cohort, and compares the results statistically (something which has not previously been reported). Secondly it expands the domain of emotion recognition to include the more abstract concept of music emotion recognition, which has shown to be impaired in other neurodegenerative diseases and to rely on similar underlying brain regions.

Finally Part 3 appraises the work presented in Parts 1 and 2, firstly by expanding on methodological limitations, and recommendations for future research, and secondly

by reflecting on the challenges of working with people suffering from a neurodegenerative disease.

TABLE OF CONTENTS

OVERVIEW **2**

TABLE OF CONTENTS **4**

LIST OF TABLES **9**

LIST OF FIGURES **10**

ACKNOWLEDGEMENTS **11**

PART 1: LITERATURE REVIEW **12**

ABSTRACT **13**

1 INTRODUCTION **14**

2 METHODS **16**

2.1 CRITERIA FOR CONSIDERING STUDIES FOR THIS REVIEW 16

2.1.1 Types of studies 16

2.1.2 Types of participants 17

2.1.3 Types of measures 18

2.2 SEARCH METHODS FOR IDENTIFICATION OF STUDIES 18

2.2.1 Electronic searches 18

2.2.2 Searching other resources 19

2.3 DATA COLLECTION AND ANALYSIS 19

2.3.1 Selection of studies 19

2.3.2 Data extraction and management 19

2.3.3 Technical assessment 20

2.3.4 Data synthesis 20

3 RESULTS **21**

3.1 DESCRIPTION OF STUDIES 21

3.1.1 Results of the search 21

3.1.2 Included studies 21

3.1.2.1 Studies of facial emotion recognition 21

3.1.2.2 Studies of auditory emotion recognition 22

3.1.2.3 Studies of emotion recognition in other modalities 23
3.1.2.4 Response options 23
3.1.2.5 Study populations 23
3.2 ASSESSMENT OF METHODOLOGICAL FACTORS 30
3.2.1 Reporting of key demographic data and results 30
3.2.2 Choice of control group 32
3.2.3 Stimulus type, presentation and response options 32
3.2.4 Statistical analysis 33
3.2.4.1 Power analysis and sample size 33
3.2.4.2 Potential confounds 34
3.2.4.3 Normality of data 35
3.2.4.4 Multiple statistical comparisons 36
3.2.4.5 Reporting of analysis and results 36
3.3 OUTCOMES 38
3.3.1 Facial emotion recognition 38
3.3.2 Vocal emotion recognition 42
3.3.3 Recognition of emotion in other modalities 42
3.3.4 Cross-modal comparisons 44
3.3.5 Disproportionate deficits in recognition of specific emotions 44
3.3.6 Within-modality stimulus type comparisons 46
3.3.7 Summary 46

4 DISCUSSION 47

4.1 METHODOLOGICAL ISSUES 47
4.1.1 Reporting demographic data and results 47
4.1.2 Choice of control group 48
4.1.3 Stimulus type, presentation and response options 49
4.1.4 Statistical analysis 50
4.1.5 Summary 52
4.2 CONCLUSIONS 53
4.2.1 Facial emotion recognition 53
4.2.2 Vocal emotion recognition 54
4.2.3 Recognition of emotion in other modalities 55

4.3 FUTURE RESEARCH 56
4.4 CLINICAL IMPLICATIONS 57
4.5 LIMITATIONS 58

5 CONCLUSION 58

6 REFERENCES 60

PART 2: EMPIRICAL PAPER 66

ABSTRACT 67

1 INTRODUCTION 69

2 METHODS 74
2.1 PARTICIPANTS 74
2.2 SAMPLE SIZE 75
2.3 MEASURES 76
2.3.1 Estimated-premorbid IQ 76
2.3.2 Executive function 77
2.3.3 Facial recognition 77
2.3.4 Musical perception 77
2.3.5 Musical background 77
2.3.6 Emotion recognition 78
2.3.6.1 Stimuli: music 78
2.3.6.2 Stimuli: facial expressions 78
2.3.6.3 Stimuli: nonverbal vocal sounds 79
2.3.6.4 General testing procedure 79
2.4 PROCEDURES 80
2.5 STATISTICAL ANALYSIS 80
2.5.1 Demographic and background data 80
2.5.2 Emotion recognition data 81
2.5.2.1 Differences between groups 81
2.5.2.2 Modality- and emotion-dependent differences in the HD group 82

3 RESULTS 83
3.1 DEMOGRAPHIC DATA 83

3.2 BACKGROUND DATA 84
3.3 EMOTION RECOGNITION 85
3.3.1 Faces 85
3.3.2 Voices 85
3.3.3 Music 85
3.3.4 Control performance 88
3.3.5 Within-group differences between emotions in each modality 88
3.3.6 Within-group differences between modalities for each emotion 88
3.3.7 Repeating analysis without participants who failed Peretz test 89

4 DISCUSSION 89
4.1 MAIN FINDINGS 89
4.2 LIMITATIONS 93
4.3 FUTURE WORK 96
4.4 CONCLUSION 98

5 REFERENCES 100

PART 3: CRITICAL APPRAISAL 108

1 INTRODUCTION 109

2 LITERATURE REVIEW 109

3 EMPIRICAL STUDY 110
3.1 RECRUITMENT 110
3.2 METHODOLOGICAL CHOICES 111
3.2.1 Executive function 111
3.2.2 Musical perception 112
3.2.3 Musical emotion recognition 112
3.2.4 Participants 114
3.2.5 Including imaging data 115
3.3 WORKING WITH PEOPLE WITH A NEURODEGENERATIVE DISEASE 116

4 SUMMARY 118

5 REFERENCES 120

APPENDIX 1: ARTICLES THAT DID NOT MEET SEARCH CRITERIA FOR LITERATURE REVIEW 121

APPENDIX 2: INDIVIDUAL STUDIES (FROM INCLUDED REFERENCES) THAT DID NOT MEET SEARCH CRITERIA FOR LITERATURE REVIEW 127

APPENDIX 3: ETHICAL APPROVAL AND RELATED PAPERWORK..... 128

APPENDIX 4: HAILSTONE MUSIC QUESTIONNAIRE..... 144

APPENDIX 5: MUSICAL BACKGROUND OF PARTICIPANTS 145

APPENDIX 6: PILOT STUDY OF MUSIC EMOTION STIMULI..... 146

APPENDIX 7: UNADJUSTED BEHAVIOURAL RESULTS FROM PART 2 .. 148

APPENDIX 8: DETAILS OF REGRESSION MODELS USED IN PART 2 .. 150

APPENDIX 9: BATTERY ORDER 154

LIST OF TABLES

Table 3-1 Characteristics of included studies 24
Table 3-2 Technical assessment 37
Table 3-3 Facial emotion recognition results 40
Table 3-4 Vocal emotion recognition results 43
Table 3-5 Emotion recognition in other modalities 43
Table 3-1 Demographic data 83
Table 3-2 Odds of a correct response (95% confidence intervals) relative to controls for each emotion in each modality, after adjusting for age, estimated pre-morbid IQ, Benton facial recognition score, Trail-Making B-A, and gender 87

LIST OF FIGURES

Figure 3.1 Mean proportion correct for each group, emotion and modality, adjusted for age, IQ, Benton, Trail-Making B-A score and gender, with 95% CIs 86

ACKNOWLEDGEMENTS

Whilst recruiting for this study I received an answer phone message from a potential participant. Without yet knowing the full details of what the study involved he said that it would be a "privilege" for him to take part in the study, and that he wanted to do all that he could to help anyone with HD. His attitude is not unique. I am constantly overwhelmed by the spirit and generosity of all my participants, HD gene carriers and their families, carers and friends. The privilege is entirely mine and my thanks goes to all of you for helping with this study and others like it.

Many others have helped me with the work presented here and I am very grateful to them all: Aakta Patel, Marianne Novak, Maggie Burrows, Miranda Say, Joy Read and Nayana Lahiri were part of the clinical team making an initial approach to potential participants and passing their details on to me - Marianne also deserves a special mention for independently going through 1724 article titles to check them against inclusion criteria for the literature review; the Track-HD admin staff and the ever-growing imaging team let me share their resources and gave me a desk and plenty of cake; Julia Hailstone kindly let me use the questionnaire she developed as part of her PhD work; Chris Frost yet again patiently advised on all things statistical; Sarah Tabrizi allowed me to work with her patients again; and special thanks goes to John King and Jason Warren, my supervisors, who guided me through the data collection and writing and have been on hand to keep me on track throughout.

Thanks also to my fellow trainees and my family; and my wonderful Steve who has supported me through one of these already, and still wanted to become my husband half-way through this one.

PART 1: LITERATURE REVIEW

EMOTION RECOGNITION IN HUNTINGTON'S DISEASE: A SYSTEMATIC REVIEW

ABSTRACT

Aims

There is increasing interest in the nature of the emotion recognition deficit in Huntington's disease (HD) with conflicting reports of disproportionate impairments for some emotions in some modalities. This review aimed to clarify the pattern of emotion recognition deficits in HD.

Methods

A systematic review and narrative synthesis was conducted for studies investigating emotion recognition in Huntington's disease. Embase, MEDLINE, PsychINFO and PubMed were searched from 1993 to 2010, and citation and reference list searches were also conducted. 1724 citations were identified.

Results

Sixteen studies met inclusion criteria. In manifest HD recognition of facial anger was found most consistently, although recognition of all negative emotions (facial and vocal) tended to be impaired. In premanifest HD impairments were inconsistent, but are seen in all facial expressions of negative emotion. Inconsistency may represent the variability inherent in HD although may also be due to between-study differences in methodology.

Conclusions

Current evidence supports the conclusion that recognition of all negative emotions tends to be impaired in HD, particularly in the facial domain. Future work should focus on using more ecologically-valid tests, and testing inter-modality differences.

1 INTRODUCTION

Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder caused by an expanded CAG repeat on chromosome 4. It is classically characterised by involuntary movements and cognitive and psychiatric deficits, with the onset of motor signs usually occurring in mid-adulthood. There is, however, often evidence for subtle cognitive and behavioural deficits ahead of these motor features (Lawrence et al., 1998; Snowden, Craufurd, Thompson, & Neary, 2002).

The ability to recognise emotions in others is a key social skill, and much work has focused on studying the expression and recognition of canonical emotions (happiness, sadness, fear, surprise, anger and disgust) which appear to be cross-cultural, and which it is argued have a biological basis (Ekman, 1992; Ekman, 1993). More recently, attempts have been made to elucidate the neural substrates of emotion recognition, with lesion studies and functional brain imaging providing new insights into the pathways underlying this skill (e.g., Calder, Keane, Lawrence, & Manes, 2004; Phillips et al., 1998; Scott et al., 1997); for a review see Adolphs (2002).

Over the last two decades there has been increasing interest in emotion recognition deficits in people with HD, and this interest is justified on both clinical and neurobiological grounds. HD is relatively common (Harper, 2002; Novak & Tabrizi, 2010) and has the potential for presymptomatic diagnosis (and therefore early intervention with disease-modifying therapies). These factors lend particular urgency to the search for biomarkers of brain dysfunction in HD, and emotion recognition is a promising candidate (Paulsen et al., 2006; Stout et al., 2011; Tabrizi et al., 2009). In addition, HD has contributed to the literature on models of emotion

recognition, as associations are made between the behavioural deficits seen in the disease, and the affected brain regions.

An initial finding of disproportionately impaired recognition of facial expressions of disgust was found in people with early symptoms of the disease (Sprengelmeyer et al., 1996). This was followed by suggestions that facial disgust recognition might also be affected in premanifest gene carriers (Gray, Young, Barker, Curtis, & Gibson, 1997), and that recognition of disgust in other modalities, such as voices, taste and odours (Hayes, Stevenson, & Coltheart, 2007; Mitchell, Heims, Neville, & Rickards, 2005), was also impaired. However other work has failed to replicate the disproportionate impairment in disgust recognition, instead suggesting that recognition of all negative emotions is broadly affected in HD (Henley et al., 2008; Johnson et al., 2007; Milders, Crawford, Lamb, & Simpson, 2003). The aim of this review was therefore to disambiguate the pattern of emotion recognition deficits in HD through a systematic appraisal of previous reports.

This review is warranted to better understand the nature and progression of cognitive impairment in HD, and how it might impact on people with HD and their carers. A better conceptualisation of the emotion recognition deficits in HD is important to improve understanding of the social interaction problems that occur in the disease. From this, more refined strategies for managing these problems might be developed. More fundamentally, the emotion processing deficit in HD and its brain mechanism may hold important clues to the pathophysiology of the disease. In particular, a selective deficit of emotion comprehension would (if substantiated) potentially predict a relatively specific pathophysiological and anatomical substrate which could

in turn be targeted as a biomarker of disease modification in future therapeutic trials (Henley, Bates, & Tabrizi, 2005; Tabrizi et al., 2009).

The purpose of this review was to appraise systematically the reported impairments in emotion recognition in HD. It aims to assess the nature of the deficits reported, and to investigate whether differences in findings can be explained by disease-related factors, such as stage or CAG repeat length. It asks what conclusions can be drawn from the current literature about which emotions people with HD struggle to recognise, and in which modalities, as well as trying to identify areas for future research.

2 METHODS

2.1 CRITERIA FOR CONSIDERING STUDIES FOR THIS REVIEW

2.1.1 Types of studies

In order to be eligible for inclusion in the review studies had to compare emotion recognition in a group of participants with Huntington's disease with a control group (i.e. quasi-experimental design). Emotion recognition was defined as any task in which stimuli conveying emotional information were presented, and for which participants were asked to state or choose which emotion they thought was represented by the stimuli. The stimulus could be of any modality (e.g. visual, vocal) and of any form (e.g. static faces, videos). Any target emotions were considered. Studies in which participants were asked to match emotions within a modality were excluded (e.g. selecting a happy face in response to a happy face stimulus); success on this task might be achieved using perceptual features alone. Studies in which

participants were asked to match emotions across a modality (e.g. selecting an angry face in response to an angry voice) were included as this cannot be solved purely on the basis of the perceptual features of the stimuli.

Studies looking exclusively at "mood" or emotion production, or semantic knowledge (e.g. about situations that might be expected to induce emotions) were excluded. Editorials, reviews, commentaries, letters or other articles that contained no original data were excluded.

2.1.2 Types of participants

The patient group had to consist of participants carrying the gene coding for Huntington's disease (Huntington's Disease Collaborative Research Group, 1993), confirmed by genetic analysis. This excluded any studies done before genetic testing was available, but ensured that findings were specific to this population.

Studies of both manifest and premanifest participants were included. Manifest HD is conventionally defined as the point at which gene carriers develop hard motor signs. Clinically, this can be a useful way in which to define disease onset, although in practice more subtle motor, cognitive and behavioural deficits are usually present many years before this point (Huntington Study Group, 1996; Paulsen et al., 2008). The control group had to be neurologically normal participants. Studies of any participants aged 18 or over were included. Participants with onset prior to this age are likely to have very high CAG repeat lengths and a rapidly progressing disease process as well as immature emotion processing mechanisms, which may be

qualitatively dissimilar to those in adults (Gao & Maurer, 2010; Kremer, 2002, pp. 43-44).

2.1.3 Types of measures

Studies must have reported a quantitative measure of emotion recognition.

2.2 SEARCH METHODS FOR IDENTIFICATION OF STUDIES

2.2.1 Electronic searches

Searches were run in the following databases: Embase (1993 to July 2010), MEDLINE (1993 to July 2010), PsychINFO (1993 to July 2010) and PubMed (1993 to July 2010).

Searches were limited from 1993 to the present day, as studies carried out prior to this would necessarily have included participants without genetically-confirmed HD.

The search used keywords "(Huntington* AND (emotion* OR cogniti* OR neuropsycho*)) NOT (mouse OR rat OR mice)". Using Ovid the search was run on Embase, MEDLINE and PsychINFO simultaneously and results were then deduplicated (a function within Ovid).

Only peer-reviewed published articles were accepted for inclusion in the review. Attempts were made to contact corresponding authors of all articles included in the review, either to ask for access to more demographic or experimental data, or to check queries about the study.

2.2.2 Searching other resources

For each study included in the review, manual searches of reference lists were conducted and a citation search was also conducted to identify further potential studies.

2.3 DATA COLLECTION AND ANALYSIS

2.3.1 Selection of studies

The initial searches identified 1724 citations (after de-duplication). The title and abstract of each citation were examined independently by both the author and a colleague, Dr Marianne Novak, against the pre-specified inclusion and exclusion criteria listed above.

104 citations could not be excluded on the basis of the title and abstract alone (a proportion of these did not have an abstract available). The full text of these citations was obtained by the author to assess whether they fully met inclusion criteria. One additional citation was identified from the reference list and citation search.

2.3.2 Data extraction and management

Data were extracted to a standardised data collection form. This covered demographic information, details of emotion recognition tests, any background tests, results, and technical assessment.

2.3.3 Technical assessment

Study structure and technical characteristics were assessed according to a number of criteria: sample size and power analysis; the nature of the control group; reporting of demographic data; the nature of stimuli, stimulus presentation and response options; ways in which potential confounding variables were measured and addressed; appropriate statistics; and reporting of quantitative outcome data. Demographic data considered necessary in order to be able to compare groups adequately between studies were: age, gender, some measure of estimated IQ or educational level, and additionally in the gene-positive group, CAG repeat length, and some estimate of disease course e.g. disease duration, United Huntington's Disease Rating Scale (UHDRS) motor score or an estimate of time to motor onset in premanifest subjects (e.g., Langbehn, Brinkman, Falush, Paulsen, & Hayden, 2004). (Note that these demographic data were considered desirable in order to assess studies fully, but these were not criteria for inclusion in the review overall).

2.3.4 Data synthesis

Given that the data reviewed here were quantitative a meta-analytic approach was considered. However ultimately a narrative synthesis was undertaken for two reasons. Firstly, although attempts were made to contact representatives of all the studies included in the review, some authors were uncontactable and this meant that quantitative results were not available for all studies. Secondly, the ways in which the HD cohorts varied between studies were not always clear (e.g. measures such as IQ, CAG repeat length, disease severity and duration were not always reported).

This meant that it would not be possible to determine the extent to which differences in effect size were attributable to these differences in the cohorts studied.

3 RESULTS

3.1 DESCRIPTION OF STUDIES

3.1.1 Results of the search

Sixteen reports met full inclusion criteria. Appendix 1 summarises reasons for which 89 studies that met initial inclusion criteria were excluded after the full text was examined. Of the sixteen studies included, one or more individual experiments from five of them were subsequently excluded for not meeting criteria (see Appendix 2).

3.1.2 Included studies

See Table 3-1 for characteristics of included studies.

3.1.2.1 Studies of facial emotion recognition

The majority of studies (14/16) included at least one test of facial emotion recognition (exceptions were Hayes et al. (2007) and Mitchell et al. (2005)). The two most commonly used face stimulus sets were 60 (or 24) static black and white images, from the Ekman and Friesen battery (Ekman & Friesen, 1976), 10 (or four) each for happiness, sadness, surprise, disgust, anger and fear, henceforth "Ekman Faces"; and the set of 30 face images[1] morphing between these six canonical emotions from the FEEST (Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002).

[1] These images were based on the face known as "JJ" from the Ekman and Friesen set

In this latter test, the "Emotion hexagon", presentation of the 30 face stimuli is repeated over six blocks, with results from the first block subsequently discarded as practice trials. Three studies also included neutral face stimuli in at least one test (Johnson et al., 2007; Snowden et al., 2008; Tabrizi et al., 2009).

Two studies used the same brief (24-stimulus) version of Ekman Faces (Gray et al., 1997; Henley et al., 2008). Some researchers opted to replace or supplement these tests with their own stimuli in a similar format (e.g., Aviezer et al., 2009; Snowden et al., 2008). One study used colour video clips made by morphing between neutral and emotional still photographs (Montagne et al., 2006).

One study also included a test of emotion recognition from eye regions only (Snowden et al., 2008).

3.1.2.2 Studies of auditory emotion recognition

Five studies included tests of emotion recognition in the auditory modality, testing either short non-verbal vocal sounds (e.g. laughter, growls) or prosody of spoken phrases constructed from non-words. Three of these studies (Calder et al., 2010; Snowden et al., 2008; Sprengelmeyer, Schroeder, Young, & Epplen, 2006) used stimuli taken from the same set of non-verbal vocal sounds (Scott et al., 1997). One study used their own non-verbal stimuli (Hayes et al., 2007), and two studies used the same "nonsense" word prosody recognition task (Sprengelmeyer et al., 1996; Sprengelmeyer et al., 2006).

3.1.2.3 Studies of emotion recognition in other modalities

One study included a test of emotion recognition of static black and white pictures portraying body language (Aviezer et al., 2009) and one, looking specifically at disgust recognition, tested recognition of pleasant and disgusting tastes and odours (Mitchell et al., 2005).

3.1.2.4 Response options

All the studies, with the exception of Mitchell et al. (2005, investigating taste and odour perception) used a forced choice response paradigm, in which participants were given a limited set of written verbal emotion terms and asked to pick the one that best described the stimulus.

3.1.2.5 Study populations

Eight studies included a sample of premanifest HD gene carriers, and 10 included a sample of people with manifest HD (i.e. unequivocal motor signs). One study included both premanifest and manifest participants in the patient group (Aviezer et al., 2009).

Study populations were drawn from a range of countries, including the United Kingdom (7), Australia (4), Germany (3), Canada (3), The Netherlands (2), France (1) and the United States (1). Ethnicities of participants in each country were not reported in any study. Culturally, this is a relatively restricted sample, based almost entirely on Western / European (the majority English-speaking) countries.

Table 3-1 Characteristics of included studies

Study	Demographic data			Stimuli & response options
		Control	*HD (premanifest & mild)*	
Aviezer et al. (2009) Expt. 1[a]	N	27	21	1) 6 b&w pictures of body language, shown 3 times on computer; 6AFC, no time limit
	Age	49.2 (10)	48.3 (10.1)	
	Gender (% F)	56%	57%	2) 40 Ekman faces, shown once on computer; 6AFC, no time limit
	CAG		42.85 (3.69)	
	Estimated IQ	-	-	
	UHDRS Motor		8.26 (8.02)	
	Country		Canada	
		Control	*Manifest HD*	
Calder et al. (2010) Study 1[b]	N	Varies from 20 to 52	21 (20 for morphs)	1) 60 Ekman faces, shown on computer for up to 3 sec; 6AFC, no time limit
	Age	"Matched"	50.43 (8.7)	
	Gender (% F)	Varies	43%	2) 30 b&w morphs "Emotion Hexagon"; 5 blocks of 30 (plus practice block), each morph shown on computer for 5 sec; 6AFC, no time limit
	CAG		Genetically confirmed in most participants	
	Estimated IQ	"Matched"	107.38 (8.40)	3) 60 non-verbal vocal sounds; 6AFC, no time limit
	UHDRS Motor		30.45 (13.10)	
	Country		United Kingdom	

Study	Demographic data			Stimuli & response options
		Control	*Premanifest HD*	
Gray et al. (1997)	N	23	17 (2 early manifest)	1) 24 Ekman faces, shown on card after 6 practice items; 6AFC, no time limit
	Age	38.26 (11.82)	38.53 (11.24)	
	Gender (% F)	-	-	
	CAG		-	
	Estimated IQ	-	-	
	UHDRS Motor		-	
	Country		United Kingdom	
		Control	*Manifest HD*	
Hayes et al. (2007)	N	14	14	1) 40 non-verbal vocal sounds; 4AFC, no time limit
	Age	51.3 (9.25)	54.6 (11.16)	
	Gender (% F)	43%	43%	
	CAG		-	
	Years in education	11.8 (2.04)	11.8 (2.12)	
	Disease duration		6.7 (5.21)	
	Country		Australia	
		Control	*Manifest HD*	
Hayes et al. (2009)	N	14	14	1) Emotion Hexagon (see Calder et al. entry, above)
	Age	51.8 (8.37)	54.6 (11.17)	
	Gender (% F)	50%	43%	2) 35 b&w morphs based on Ekman faces, at different intensities ranging from 0 to 150%, 5 blocks of 35 (plus 5 practice stimuli), each morph shown on computer; 6AFC
	CAG		-	
	Years in education	11.8 (1.81)	11.8 (2.12)	
	Disease duration		6.7 (5.21)	
	Country		Australia	

Study	Demographic data				Stimuli & response options
		Control	*Premanifest*	*Manifest*	
Henley et al. (2008)	N	20	21	40	1) 24 Ekman faces, shown on card after 6 practice items; 6AFC, no time limit
	Age	44.9 (10.5)	37.2 (7.9)	48.5 (9.6)	
	Gender (% F)	65%	52%	50%	
	CAG		42.2 (1.8)	43.7 (2.4)	
	Estimated IQ	106.2 (11.6)	103.2 (9.3)	105.3 (13.0)	
	UHDRS Motor	1.1 (0.9)	3.6 (4.0)	28.9 (12.6)	
	Country		United Kingdom		
		Control	*Premanifest HD*		
Hennenlotter et al. (2004)	N	9	9		1) Emotion Hexagon (see above)
	Age	"Matched"	37.4 (5.4)		
	Gender (% F)	44%	44%		
	CAG		43.7 (1.7)		
	Estimated IQ	"Matched"	112.9 (11.1)		
	UHDRS Motor				
	Country		Germany		
		Control	*Premanifest HD*		
Johnson et al. (2007)	N	57	464		1) 70 Ekman faces, shown on computer touch screen for up to 4 sec after 7 practice trials using verbal labels instead of faces; 7AFC, up to 8 sec to respond using touch screen
	Age, yr	43.01 (10.13)	41.43 (9.63)		
	Gender (%F)	61%	63%		
	CAG	<30	>39		
	Years in education	15.11 (2.29)	14.48 (2.59)		
	UHDRS Motor				
	Country		United States, Canada, Australia		

Study	Demographic data				Stimuli & response options
		Control	*Premanifest HD*		
Kipps et al. (2007)	N	13	17		1) Emotion Hexagon (see above)
	Age, yr	42.0 (11.4)	43.8 (10.0)		
	Gender (%F)	31%	47%		
	CAG	20 (3.3)	41 (2.8)		
	Estimated IQ	-	-		
	UHDRS Motor	3.6 (1.8)	6.4 (3.9)		
	Country		Australia		
		Control	*Premanifest*	*Manifest*	
Milders et al. (2003)	N	20	20	20	1) 60 Ekman faces, shown on card; 6AFC, no time limit
	Age, yr	47.9 (9.3)	38.4 (9.5)	47.6 (8.45)	
	Gender (%F)	40%	65%	40%	
	CAG		-	-	
	Estimated IQ	109.0 (6.0)	110.1 (6.1)	105.8 (7.41)	
	Disease duration			6.5 (3.2)	
	Country		United Kingdom		
		Control	*Manifest HD*		
Mitchell et al. (2005)[c]	N	8 (6 for odours)	8 (6 for odours)		1) 5 disgusting & 5 pleasant odours, presented once for up to 5 sec; rate odour on 10cm anchored line scale from very pleasant to very disgusting
	Age, yr	49.25 (4.86)	53.25 (7.25)		
	Gender (%F)	50%	50%		
	CAG		-		2) 6 everyday foods, presented individually and then in 4 appropriate & 4 inappropriate pairings; rate taste as for expt. 1 above
	Estimated IQ	-	-		
	UHDRS Motor		-		
	Country		United Kingdom		

Study	Demographic data			Stimuli & response options
		Control	*Manifest HD*	
Montagne et al. (2006)	N	30	7	1) 54 colour videos (9 different intensities for each of 6 emotions), made by morphing stills from actors, presented in 9 blocks of increasing intensity (20% - 100%); 6AFC, no time limit
	Age, yr	39.0 (11.1)	46.4 (11.2)	
	Gender (%F)	53%	29%	
	CAG		-	
	Estimated IQ	-	-	
	UHDRS Motor		17.1 (6.2)	
	Country		The Netherlands	
		Control	*Manifest HD*	
Snowden et al. (2008)[d]	N	12	10	1) 60 Ekman faces; 6AFC, no time limit
	Age, yr	57 (9)	47 (9)	
	Gender (%F)	33%	50%	2) 60 Ekman faces; 2AFC, no time limit
	CAG		-	
	Estimated IQ	108.0 (6.7)	103.6 (10.7)	3) 120 non-verbal vocal sounds (20 for each of 6 emotions); 6AFC, no time limit
	Disease duration		7 (3)	
	Country		United Kingdom	4) 35 b&w faces, "Manchester" set (5 for each of 7 emotions plus 5 practice items); 7AFC, no time limit
				5) 35 b&w eye regions, "Manchester" set (as (4)); 7AFC, no time limit
		Control	*Manifest HD*	
Sprengelmeyer et al. (1996)	N	17	13 (11 for study 2)	1) Emotion Hexagon (see above)
	Age, yr	50.7 (14.3)	45.0 (7.6)	
	Gender (%F)	47%	54%	2) 60 Ekman faces, shown on computer for up to 3 sec; 6AFC, no time limit
	CAG		45.2 (4.9) (N=11)	
	Estimated IQ	107.5 (10.0)	105.6 (10.7)	3) 60 "nonsense" sentences spoken with emotional prosody; 6AFC, no time limit
	Disease duration		6.6 (2.5)	
	Country		Germany	

Study	Demographic data				Stimuli & response options
		Control	*Premanifest HD*		
Sprengelmeyer et al. (2006)	N	8 (6)[e]	14 (12)[e]		1) 60 Ekman faces, shown on computer for up to 3 sec; 6AFC, no time limit
	Age, yr	38.3 (14.5)	31.0 (8.5)		
	Gender (F:M)	75%	64%		2) Emotion Hexagon (see above)
	CAG	20.4 (3.8)	45.1 (4.0)		
	Estimated IQ	108.8 (9.7)	113.2 (8.1)		3) 60 "nonsense" sentences spoken with emotional prosody; 6AFC, no time limit
	UHDRS Motor				
	Country		Germany		4) 60 non-verbal vocal sounds; 6AFC, no time limit
		Control	*Premanifest*	*Manifest*	
Tabrizi et al. (2009)[f]	N	123	120	123	1) 70 Ekman faces, shown on computer touch screen for up to 4 sec after 7 practice trials using verbal labels instead of faces; 7AFC, up to 8 sec to respond using touch screen
	Age, yr	46.1 (10.2)	40.8 (8.9)	48.8 (9.9)	
	Gender (%F)	55%	55%	54%	
	CAG		43.1 (2.4)	43.7 (3.0)	
	Years in education[g]	4.0 (1.3)	3.9 (1.2)	3.6 (1.3)	
	UHDRS Motor			23.7 (10.8)	
	Country		Canada, France, The Netherlands, United Kingdom		

Note: "Ekman faces" are always black & white (see text)
A hyphen denotes data not reported; an empty cell denotes variable not applicable
AFC = Alternative Forced Choice
Attempts were made to contact representatives of all studies included in the review to resolve queries; responses were received from most authors other than Drs Hayes, Hennenlotter and Sprengelmeyer
[a] CAG and UHDRS motor data provided by Dr Hillel Aviezer (personal communication)
[b] Discrepant age data and CAG confirmation provided by Prof Andy Calder (personal communication)
[c] Age and gender data provided by Dr Ian Mitchell (personal communication)
[d] IQ data provided by Dr Julie Snowden and Dr Jennifer Thompson (personal communication)
[e] Figures in brackets denote N at timepoints 2 and 3; data are given for all subjects at timepoint 1
[f] CAG repeat length and UHDRS motor score provided by Prof Sarah Tabrizi and the Track-HD team (personal communication)
[g] UNESCO International Standard Classification of Education: Level 3 qualifications typically start at the end of compulsory education (after 11 years of schooling in the UK)

3.2 ASSESSMENT OF METHODOLOGICAL FACTORS

Study methodology was assessed in four distinct categories: choice of control group; reporting of key demographic data and results; stimulus type, presentation and response options; and reporting of statistical analysis, including discussion of power, potential confounds and the issue of multiple comparisons (Table 3-2).

3.2.1 Reporting of key demographic data and results

Huntington's disease is highly heterogeneous, and clinical presentation is known to depend on age and CAG repeat length (and their interaction), which explain some of the variance in age of motor onset (see e.g., Mahant, McCusker, Byth, & Graham, 2003; Rosenblatt et al., 2006). It is therefore important to be able to rule out differences in the clinical characteristics of cohorts as a potential cause of differences between study findings. In addition, factors such as age, education and intelligence, and possibly gender, may affect performance on cognitive tasks in both HD and control groups. The impact of these factors both within studies (between patient and control groups) and between studies needs to be taken into account when assessing differences in outcome. Consequently it is important for studies to report summary data for each of these variables, so that the effects (if any) of these potential confounds can be judged.

Four studies were considered to have reported adequate demographic data: age, gender, an index of intellectual ability, CAG repeat length, and an index of disease severity (Henley et al., 2008; Hennenlotter et al., 2004; Sprengelmeyer et al., 1996; Sprengelmeyer et al., 2006). Most others reported most of the above variables but many did not have CAG repeat data available; in these cases although participants

had undergone genetic testing for confirmation of diagnosis, researchers had not always requested (or been granted access to) the exact CAG data. Some lacked an estimate of intelligence or educational level, although some authors were able to provide the extra data on request.

Also of note is the fact that studies varied in their definition of pathological CAG repeat length. Typically alleles of up to 35 repeats are considered normal, whilst alleles with 40 or more repeats are fully penetrant and the carrier is likely to show signs of HD within a normal lifespan. The intermediate repeat numbers (36-39) are not fully penetrant but there have been reports of 36 CAG repeats leading to the disease, and of people living into their 90s with 39 CAG repeats and no signs of HD (Rubinsztein et al., 1996). Whilst the majority of studies tend to include participants with a CAG repeat length of 40 or above, at least two included participants with CAG repeat lengths between 36 and 39 (Aviezer et al., 2009; Tabrizi et al., 2009), and not all report their criteria. This raises the possibility that some participants may not be representative of the more general HD population.

The majority of studies reported their findings in full (i.e. gave quantitative measures of central tendency and spread in each of the groups tested). A number of authors made their raw data available on request if they were not available in the published paper. Some authors preferred to report composite scores (Sprengelmeyer et al., 2006; Tabrizi et al., 2009), and were able to justify this, although this makes it difficult to draw direct comparisons between individual tests.

3.2.2 Choice of control group

Control groups were of three different kinds. Ten studies used healthy volunteers as controls; five used gene-negative controls from an HD environment (either people who had been at risk and tested negative for the HD gene, or partners of gene-positive participants); and one study used at-risk gene-negative controls who were unaware of their negative gene status when they completed the study tests.

3.2.3 Stimulus type, presentation and response options

As mentioned above, the majority of studies used very similar stimulus sets. Facial stimuli were usually based on the Ekman and Friesen set (Ekman & Friesen, 1976). However there are a number of variations of this: whether simple faces or the Emotion Hexagon are used, whether or not neutral faces are included, and the overall number of stimuli used. Only two studies included non-Ekman facial stimuli, one using a similar black and white static set (Snowden et al., 2008) and one making their own colour videos from actors (Montagne et al., 2006).

Similarly, studies of vocal emotion recognition tended to use stimuli drawn from the same set (either the non-verbal vocal sounds, see Calder et al., 2004; Scott et al., 1997; or the prosodic stimuli used by Sprengelmeyer et al., 1996). One exception was Hayes et al. (2007) where it seems that novel non-verbal vocal stimuli were used.

There were a number of subtle variations of presentation and response options, particularly pertinent to facial stimuli as these are not naturally time-limited (as for example are auditory stimuli). Some faces were presented for a limited time, others

were presented until participants had made a choice. As mentioned above, in the Emotion Hexagon test each stimulus is repeated six times (responses from the first presentation being discarded as practice items), whereas in simple face tests each stimulus is presented only once.

Most studies did not impose a fixed time to respond, but two limited the time available in which a response could be made (Johnson et al., 2007; Tabrizi et al., 2009). All studies of facial and vocal emotion recognition used an alternative forced choice (AFC) response paradigm, but they varied in the choices given. Most gave the same number of response options as there were emotion categories (e.g. if six different emotions were presented, there would be six response options). One deliberately reduced the number of responses (Snowden et al., 2008) in order to evaluate performance when task demands were decreased. One study used a 6AFC response when only four different emotions were represented in the stimuli (Aviezer et al., 2009).

3.2.4 Statistical analysis

3.2.4.1 Power analysis and sample size

Only two studies reported considerations of power and sample size calculation (Johnson et al., 2007; Tabrizi et al., 2009). These were the two largest studies (over 100 participants in the patient groups), and both had calculated that they were adequately powered to detect relatively small effects in premanifest and manifest HD populations. Sample sizes in the remaining studies ranged from six participants with manifest HD (Mitchell et al., 2005, odour test) to 40 (Henley et al., 2008).

3.2.4.2 Potential confounds

Potential confounds were dealt with in a number of different ways. Almost all studies reported gender, age and some measure of estimated pre-morbid intelligence or educational level. Most of these reported that groups were "matched" for one or more of these variables, and some, but not all, reported a statistic to confirm that there were no statistically significant differences between groups. Four studies included some or all of these variables as covariates in their analysis (Henley et al., 2008; Johnson et al., 2007; Milders et al., 2003; Tabrizi et al., 2009).

Almost all studies that tested facial emotion recognition included a standard test of face recognition[1], and sometimes a test of visual acuity or contrast sensitivity, and other visual or face processing tasks (the exception was Tabrizi et al. (2009), which was designed to assess potential biomarkers, rather than to test facial emotion recognition *per se*). One study excluded two participants with poor acuity (Aviezer et al., 2009), and no other studies reported impairments in basic visual skills. Six studies reported that performance on the Benton Facial Recognition Test was significantly worse in HD groups than control groups. Two took this into account in their analysis: one adjusted for facial recognition ability by including Benton score as a covariate in the analysis (Henley et al., 2008) and one investigated whether Benton scores correlated with emotion recognition scores (Snowden et al., 2008). The others did not take poor Benton performance into account in the emotion recognition analysis although mean (SD) data in most cases suggest that some participants may

[1] Benton Facial Recognition Test (Benton, Sivan, Hamsher, Varney, & Spreen, 1978)

have fallen into the "moderately impaired" range. In two studies in which a group difference was not found on the Benton, some HD participants still scored in the "moderate impairment" or "severe impairment" range although this is not commented on (Aviezer et al., 2009; Gray et al., 1997). In addition to those studies that reported group differences in Benton score, one study included Benton as a covariate in the main analyses although did not report whether group differences were statistically significant (Johnson et al., 2007).

Studies that included auditory stimuli did not report testing auditory perception. The one study that investigated taste and olfactory recognition tested olfactory identification and threshold and excluded two HD participants from the olfactory experiment on the basis of poor performance (Mitchell et al., 2005).

3.2.4.3 Normality of data

The majority of studies noted that data did not meet assumptions needed for parametric statistics. Many cited ceiling effects and used non-parametric tests (e.g. techniques such as Mann-Whitney tests, or bootstrap confidence intervals), whilst some just acknowledged heterogeneity of residual variance between groups and used appropriate statistics for this.

Five studies used parametric statistics and did not discuss whether the data were normally distributed (Aviezer et al., 2009; Hayes et al., 2007; Milders et al., 2003; Mitchell et al., 2005; Montagne et al., 2006). One study used non-parametric statistics for behavioural data, but opted to use parametric statistics on untransformed

mean reaction time data, although it would seem likely that such data might have been positively skewed (Sprengelmeyer et al., 2006).

3.2.4.4 Multiple statistical comparisons

All studies reported several statistical comparisons. Three studies reported Bonferroni-corrected results (controlling the false positive rate across a number of comparisons) (Calder et al., 2010; Hayes et al., 2007; Milders et al., 2003). Three studies discussed the increased risk of false positives but preferred to maximise power by reporting uncorrected results (Henley et al., 2008; Snowden et al., 2008; Sprengelmeyer et al., 2006). The remaining ten studies did not discuss this issue.

3.2.4.5 Reporting of analysis and results

The majority of studies reported their analysis and test results clearly. Most studies reported two-tailed tests. Four studies chose to use one-tailed tests for some or all of their comparisons based on *a priori* predictions that the HD group would, on average, do worse than controls (Gray et al., 1997; Hennenlotter et al., 2004; Kipps, Duggins, McCusker, & Calder, 2007; Sprengelmeyer et al., 2006) although despite this, Sprengelmeyer et al. (2006) report one test in which the HD group outperformed the control group.

Table 3-2 Technical assessment

Study	Control type	Power analysis / sample size	Normality of data considered	Multiple comparisons addressed	Confounds measured and controlled for
Aviezer et al. 2009	■	■	■	■	■
Calder et al. 2010	■	■	■	■	■
Gray et al. 1997	■	■	■	■	■
Hayes et al. 2007	■	■	■	■	■
Hayes et al. 2009	■	■	■	■	■
Henley et al. 2008	■	■	■	■	■
Hennenlotter et al. 2004	■	■	■	■	■
Johnson et al. 2007	■	■	■	■	■
Kipps et al. 2007	■	■	■	■	■
Milders et al. 2003	■	■	■	■	■
Mitchell et al. 2005	■	■	■	■	■
Montagne et al. 2006	■	■	■	■	■
Snowden et al. 2008	■	■	■	■	■
Sprengelmeyer et al. 1996	■	■	■	■	■
Sprengelmeyer et al. 2006	■	■	■	■	■
Tabrizi et al. 2009	■	■	■	■	■

Control type: ■ = gene-negative / spouses; ■ = gene-negative unaware of status; ■ = healthy volunteers
Power analysis / sample size: ■ = analysis performed; ■ = no analysis performed
Normality of data considered: ■ = discussed, and stats adapted accordingly; ■ = stats adjusted for inhomogeneity of variance only; ■ = not discussed
Multiple comparisons addressed: ■ = discussed and addressed; ■ = not discussed
Confounds measures and controlled for: ■ = effect of potentially confounding variables taken into account in analysis; ■ = effect of some potentially confounding variables taken into account in analysis; ■ = not considered

3.3 OUTCOMES

Outcomes are described separately for manifest and premanifest populations as this is how the majority of studies were designed. As discussed above, this is a somewhat arbitrary distinction, based on an assessment of motor symptoms. There is inevitably variation between clinicians with regard to when symptoms are sufficient to make a diagnosis of manifest disease, and different studies use different cut-off points to define this. Aviezer et al. (2009) included both manifest and premanifest participants in a single group in their study. Examination of their data shows that nine out of 21 participants who completed the study had a UHDRS motor score of five or less, a cut-off used elsewhere to discriminate between manifest and premanifest participants (Tabrizi et al., 2009). Therefore since the majority of participants had clear signs of the disease, Aviezer et al.'s results are reported below as representing manifest HD.

3.3.1 Facial emotion recognition

In participants with manifest HD the most consistent impairment was shown for recognition of facial anger: this was found to be impaired in every study that included a test for it. Disgust recognition was also found to be impaired in almost every study that tested it, with the exception of Snowden et al. (2008) in a task in which they gave a two-alternative forced choice response option to the Ekman faces, instead of the usual six (i.e. they reduced task demands); anger and fear recognition were still impaired in this condition. Fear recognition was often found to be impaired, although only using Ekman stimuli (not moving facial stimuli (Montagne et al., 2006), or a non-Ekman stimulus set, the "Manchester" set, a different, locally-

made set of black-and-white static emotion faces (see Snowden et al., 2008). Sadness and surprise recognition were found to be impaired less often, whilst an impairment in happiness recognition was only found by two groups (Calder et al., 2010; Hayes, Stevenson, & Coltheart, 2009) (Table 3-3).

In premanifest participants an impairment in disgust recognition was most frequently reported and was seen in five out of eight studies. Three of these reported that disgust was selectively impaired (Gray et al., 1997; Hennenlotter et al., 2004; Sprengelmeyer et al., 2006), whilst two found an impairment across negative emotions (Johnson et al., 2007; Tabrizi et al., 2009). Sprengelmeyer et al. (2006) also reported a deficit in surprise recognition at the third of three timepoints tested. The remaining three studies found no evidence of impairment at all in premanifest participants, although one study explained a finding of impaired happiness recognition as an artefact of the ceiling effect in controls for that emotion (Henley et al., 2008).

Snowden et al. (2008) found that manifest HD participants were impaired at recognising sadness and disgust from the eye regions alone (Table 3-3).

Table 3-3 Facial emotion recognition results

Population	Study	Stimuli	Ha	Sa	Su	Di	An	Fe	Ne
Manifest HD	Aviezer et al. 2009	Ekman faces	○	○		●	●		
	Calder et al. 2010	Ekman faces	●	●	●	●	●	●	
		Emotion hexagon	●	●	●	●	●	●	
	Hayes et al. 2009	Emotion hexagon	○	●	●	●	●	●	
		Ekman faces at different intensities	○	●		●	●	●	
	Henley et al. 2008	Ekman faces	○	○	●	●	●	●	
	Milders et al. 2003	Ekman faces	○	●	○	●	●	●	
	Montagne et al. 2006	Videos at different intensities	○	○	○	●	●	○	
	Snowden et al. 2008	Ekman faces (6AFC)	○	●	●	●	●	●	
		Ekman faces (2AFC)	○	○	○	○	●	●	
		Manchester faces	○	○	○	●	●	○	○
	Sprengelmeyer et al. 1996	Emotion hexagon	○	●	●	●	●	●	
		Ekman faces	○	●	●	●	●	●	
	Tabrizi et al. 2009[a]	Ekman faces	(○)	(●)	(○)	(●)	(●)	(●)	(○)
Manifest HD	Snowden et al. 2008	Manchester eyes	○	●	○	●	○	○	○

Population	Study	Stimuli	Ha	Sa	Su	Di	An	Fe	Ne
Premanifest HD	Gray et al. 1997	Ekman faces	○	○	○	●	○	○	
	Henley et al. 2008	Ekman faces	●	○	○	○	○	○	
	Hennenlotter et al. 2004	Emotion hexagon	○	○	○	●	○	○	
	Johnson et al. 2007	Ekman faces	○	●	○	●	●	●	○
	Kipps et al. 2007	Emotion hexagon	○	○	○	○	○	○	
	Milders et al. 2003	Ekman faces	○	○	○	○	○	○	
	Sprengelmeyer et al. 2006[b]	Ekman faces + Emotion hexagon	○	○	(●)	●	○	○	
	Tabrizi et al. 2009[a]	Ekman faces	(○)	(●)	(○)	(●)	(●)	(●)	(○)

Key: Ha = happiness; Sa = sadness; Su = surprise; Di = disgust; An = anger; Fe = fear; Ne = neutral
● = group difference; ○ = no group difference; blank = not tested
[a] Composite "negative emotion" score tested
[b] Result in brackets only found at timepoint 3

3.3.2 Vocal emotion recognition

In manifest participants an impairment in vocal disgust recognition was found consistently (four out of four studies) for both short non-verbal vocal sounds and speech prosody. Anger recognition was found to be impaired in the three studies using non-verbal vocal sounds, but not for prosody. Fear recognition was also found to be impaired in three out of four studies, including both non-verbal vocal sounds and prosody. Using prosodic stimuli, impairments were also reported for recognising surprise and happiness (Sprengelmeyer et al., 1996) although no other studies reported deficits in these emotions. No studies reported a deficit of sadness recognition from vocal sounds (Table 3-4).

Only one study tested vocal emotion recognition in a premanifest cohort (Sprengelmeyer et al., 2006) and reported no evidence of impairments at any of three timepoints tested, using a combined score from sounds and prosodic stimuli (Table 3-4).

3.3.3 Recognition of emotion in other modalities

Aviezer et al. (2009) found no evidence that their HD population was impaired at recognising sad, disgusted or angry body language. Mitchell et al. (2005) reported that their manifest HD cohort tended to rate unpleasant odours and taste combinations as less disgusting than controls (Table 3-5).

Table 3-4 Vocal emotion recognition results

Population	Study	Stimuli	Ha	Sa	Su	Di	An	Fe	Ne
Manifest HD	Calder et al. 2010	Non-verbal vocal sounds	○	○	○	●	●	●	
	Hayes et al. 2007	Non-verbal vocal sounds		○		●	●	○	
	Snowden et al. 2008	Non-verbal vocal sounds	○	○	○	●	●	●	
	Sprengelmeyer et al. 1996	Prosody	●	○	●	●	○	●	
Premanifest HD	Sprengelmeyer et al. 2006	Prosody + non-verbal vocal sounds	○	○	○	○	○	○	

Key: Ha = happiness; Sa = sadness; Su = surprise; Di = disgust; An = anger; Fe = fear; Ne = neutral
● = group difference; ○ = no group difference; blank = not tested

Table 3-5 Emotion recognition in other modalities

Population	Study	Stimuli	Ha	Sa	Su	Di	An	Fe	Ne
Manifest HD	Aviezer et al. 2009	Body language		○		○	○		
	Mitchell et al. 2005	Odours				●			
		Tastes				●			

Key: Ha = happiness; Sa = sadness; Su = surprise; Di = disgust; An = anger; Fe = fear; Ne = neutral
● = group difference; ○ = no group difference; blank = not tested

3.3.4 Cross-modal comparisons

Several studies included tests of emotion recognition in more than one modality, although they only report independent statistics for group differences in each modality (i.e., they do not directly compare performance between modalities statistically). Aviezer et al. (2009) reported that anger and disgust recognition from Ekman faces was impaired in the absence of impairments in recognising emotional body language. Calder et al. (2010) found a global impairment in recognising the six canonical emotions from Ekman faces, but only disgust, anger and fear recognition were impaired from non-verbal vocal sounds. A similar pattern was shown by Snowden et al. (2008) using the same face stimuli: recognition of all emotions except happiness was impaired with facial stimuli, whilst only disgust, anger and fear were impaired with non-verbal vocal sounds. However using a different set of faces Snowden et al. (2008) reported impairment only for disgust and anger recognition. In the study of Sprengelmeyer et al. (1996) subjects showed impaired recognition of surprise, disgust and fear from both facial expressions and prosody, but impaired sadness and anger recognition only from facial expressions. In a premanifest cohort, a selective impairment recognising the facial expression of disgust was found in the absence of impairments for the other five canonical facial emotions, or any deficits in recognising emotions from sounds or prosody (Sprengelmeyer et al., 2006).

3.3.5 Disproportionate deficits in recognition of specific emotions

Early reports suggested that HD gene carriers (both manifest and premanifest) were disproportionately impaired at recognising disgust, both using faces and prosody (Gray et al., 1997; Sprengelmeyer et al., 1996). Sprengelmeyer et al. (1996) tested this statistically, converting manifest HD performance to proportion of controls (to

adjust for emotion difficulty) and comparing disgust recognition with the next worst recognised emotion, fear. Disgust recognition was significantly worse than fear recognition for both Ekman Faces, Emotion Hexagon, and prosodic stimuli; HD participants scored at or below chance level. Sprengelmeyer et al. (2006) also report that in a premanifest cohort facial disgust recognition was the only emotion impaired relative to controls, and that 5/12 gene carriers were only impaired at disgust (judged by z scores), although the other seven were either unimpaired, or globally impaired. Three other studies report selective or disproportionate impairments in disgust recognition. Gray et al. (1997) and Hennenlotter et al. (2004) both found that facial disgust recognition was the only emotion impaired in premanifest cohorts, although this was not assessed statistically in relation to other emotions. Hayes et al. (2007) reported that more HD participants had z scores of >-1.56 (and more scored at chance) for non-verbal vocal disgust recognition than for any other vocal emotion, although again, differences between emotions were not assessed statistically.

Most other studies do not report selective or disproportionate impairments in disgust recognition. Hayes et al. (2009) tested differences between emotions and found no evidence that one was more impaired than any other (Ekman Faces and Emotion Hexagon). Henley et al. (2008) compared emotion recognition performance statistically (adjusting for control scores) and found recognition of anger to be disproportionately impaired (Ekman Faces). Milders et al. (2003) found that recognition of disgust was less impaired than recognition of anger, fear and sadness (Ekman Faces). Snowden et al. (2008) reported either no impairment of disgust recognition (Ekman Faces with 2AFC), or that no patient got their worst scores at disgust recognition (Ekman Faces with 6AFC); when assessed statistically, disgust recognition in their HD group was no worse (and in one case was better) than fear

and anger recognition both for the Manchester faces and eyes set (see section 3.3.1), and for non-verbal vocal sounds. Other studies did not test inter-emotion differences statistically but argued that the pattern of findings did not support a disproportionate impairment of disgust recognition (e.g., Aviezer et al., 2009; Calder et al., 2010, who found a greater number of HD participants impaired at anger across all tasks; Johnson et al., 2007; Montagne et al., 2006).

3.3.6 Within-modality stimulus type comparisons

Many studies used both the "Ekman Faces" set, as well as the "Emotion Hexagon". Calder et al. (2010) reported similar deficits (across all emotions except happiness) with both stimulus sets. Hayes et al. (2009) found that sadness, surprise, disgust, anger and fear recognition were impaired using both the Emotion Hexagon, and Ekman Faces at varying intensities, although happiness was only impaired on the Ekman Faces set, and no impairment was seen for 25% and 50% sad Ekman Faces. Snowden et al. (2008) compared the Ekman Faces set 6AFC, with 2AFC, an alternative face set, and eye regions only. They reported a recognition deficit across all emotions except happiness using Ekman Faces, reduced to anger and fear recognition deficits when the task was simplified (2AFC), whilst only disgust and anger recognition were impaired on the alternative "Manchester" set, and sadness and disgust recognition using eye regions only.

3.3.7 Summary

Disgust, anger and fear recognition were most often impaired in manifest HD populations, across modalities. For face recognition, the most frequently tested modality with the most consistency in stimulus presentation, a deficit in anger recognition was found in all the studies in which it was tested. In premanifest

populations deficits were more commonly seen for disgust recognition than any other emotion, but only in the facial modality. Whilst disgust recognition appears disproportionately impaired in some HD populations, this is not true of all populations tested, nor across modalities. Outcome varies depending on the type of stimulus used (even within a modality).

4 DISCUSSION

This review examined 16 studies, investigating emotion recognition in manifest and premanifest HD across different modalities. The discussion will first address methodological issues, before going on to draw conclusions and discuss what future research is needed, and the clinical implications and limitations of the review.

4.1 METHODOLOGICAL ISSUES

4.1.1 Reporting demographic data and results

Very few studies reported adequate demographic data as well as quantitative results, although the studies by Calder et al. (2010), Hayes et al. (2009), Kipps et al. (2007), Milders et al. (2003) and Snowden et al. (2008) reported all their results and most of the demographic data listed above. Two studies that reported very little demographic data were those of Gray et al. (1997) and Mitchell et al. (2005), although the latter author was able to provide age and gender data on request (Mitchell, personal communication). Two studies reported only composite scores (across tasks, Sprengelmeyer et al., 2006; across emotions, Tabrizi et al., 2009) which made it hard to calculate effect sizes and compare results with other studies.

One of the biggest weaknesses of the literature is the fact that so few studies report sufficient demographic data to allow the HD cohorts to be compared. Many report

either UHDRS motor score, or disease duration, but not both, making it hard to judge cohort similarity, and very few report CAG repeat length data, which again would be useful in conceptualising how severe or advanced the cohort is likely to be. Therefore one major cause of inter-study variability, the inherent heterogeneity in HD, cannot be investigated thoroughly, although this would be a useful exercise.

4.1.2 Choice of control group

The rationale for using gene-negative or partner controls is that these people live in a similar social and emotional environment to people with HD; interacting with family members whose emotion recognition, and possibly expression, is impaired may impact on the controls' expression and recognition of emotion, and they are more likely than unrelated volunteers to be subject to similar stresses, and therefore have similar levels of anxiety and depression. Use of a gene-negative control group therefore aims to minimise group differences attributable to social or emotional factors. One study showed that gene-negative control performance fell below published norms at recognising sad, angry and fearful faces from the Ekman 60 set (Henley et al., 2008). Both Sprengelmeyer et al. (2006) and Gray et al. (1997) found that their gene-negative control groups performed worse than healthy volunteer controls at recognising anger (Ekman Faces and Emotion Hexagon stimuli), although in the latter case the additional stress of undergoing genetic testing and being unaware of the result may have contributed to poor performance. Six studies (including Gray et al. (1997) whose gene-negative participants were unaware of their status) used gene-negative controls, and 10 did not. It is acknowledged that recruiting a gene-negative control group may be more difficult than recruiting healthy volunteers. However, given the evidence cited above that otherwise healthy people living in an HD environment show emotion recognition impairments relative

to people who are not exposed to those environments, using a gene-negative control group may be preferable in order to reduce the likelihood that differences can be explained by non-organic disease effects.

4.1.3 Stimulus type, presentation and response options

Another potential difficulty in interpreting results across studies is the use of slightly different stimulus sets. The two main tools for assessing facial emotion recognition in the literature – the Emotion Hexagon and the Ekman faces – are not of equivalent difficulty (based on the mean percent correct achieved by Ekman's normative sample). In addition, the Emotion Hexagon is not a simple facial emotion recognition task. Morphs of a single subject ("JJ") were constructed by blending different proportions of two emotions, placing each next to one it was most likely to be confused with (Calder et al., 1996) although in fact, in order to fit all six canonical emotions into the hexagon this is not always true for each pair of emotions. Blocks tend to be repeated in testing which might inflate differences between controls and HD subjects if controls benefited from learning over the earlier blocks whilst HD subjects did not. However, based on the literature reviewed here, results using these two different sets of facial stimuli are fairly similar (when tested on the same cohort) suggesting that results are comparable. Additionally, limiting presentation time and response time did not seem to impact hugely on outcome.

Interestingly, in the one study that compared the Ekman Faces to a different face set, the Ekman set seemed to be harder (Snowden et al., 2008). All studies except one used static black and white faces, which may not be very ecologically valid. However the one study to use colour videos had a relatively small HD group and

groups were not well-matched for gender, and therefore these results would benefit from replication (Montagne et al., 2006).

There was also a suggestion that different deficits were seen in vocal emotion recognition depending on whether non-verbal sounds or prosodic stimuli were used. However far fewer studies have tested vocal emotion recognition (compared with face emotion recognition), and prosodic stimuli have only been used in one manifest and one premanifest cohort, so these results would also benefit from replication. Sprengelmeyer et al. (2006) stated that scores on non-verbal vocal and prosodic stimuli in their premanifest cohort were correlated, and therefore presented a composite score (using which there was no evidence of impairment). However it would be interesting to compare the two stimulus sets directly in the same cohort to begin to ascertain whether deficits vary depending on stimulus type.

4.1.4 Statistical analysis

Only two studies considered issues of power and sample size, both stating that they were powered adequately to detect relatively small effects (Johnson et al., 2007; Tabrizi et al., 2009). Sample sizes in some of the other studies were relatively small (in some cases fewer than 10 participants in the HD group) which raises the question of whether some studies, particularly of premanifest populations where effects are likely to be small, were under-powered.

Most, although not all, control and HD groups were said to be matched for demographic variables such as age, gender, and in some cases IQ or educational level. However, non-statistically significant group differences in these variables does not mean that small differences cannot still influence outcome on another variable. Johnson et al. (2007) show that age, gender and education have

independent effects on recognition of some emotions, which, on average, decreases with increasing age and less education, and is better in females than in males. There is also evidence that fear and anger recognition might deteriorate with age, whilst disgust recognition is relatively spared (Calder et al., 2003). The original finding of impaired disgust recognition in HD used controls who were on average five years older than HD subjects, and did not adjust for the effects of age (Sprengelmeyer et al., 1996), and so it may be that this group difference inflated the effect.

It is notable that although a number of studies included the Benton Facial Recognition Test as a "background measure", few took these results into account in their analysis. Several studies reported that, on average, HD groups performed worse than control groups at this test, and in many cases were in the "borderline range". In fact examination of mean (SD) face recognition scores (where available) suggests that some HD participants may even have been moderately or severely impaired (and data provided by Aviezer et al. (2009) and Gray et al. (1997) confirm this to be the case even when no statistically significant group differences were found). This means that some group differences on facial emotion recognition may in fact be attributable to poorer face processing in some HD participants.

Where auditory stimuli were used, no study reported background tests of audition. Auditory stimuli are also naturally transient, unlike visual stimuli which (in many studies) remained present until a response was made. However no study reported making allowances for the potential demands on working memory this makes by, for example, allowing participants to hear auditory stimuli more than once if needed.

Two statistical issues that arise continually in this field are those of non-normally distributed data (groups often perform at ceiling on happiness recognition, for

example) and carrying out multiple statistical comparisons. Larger, and more recent studies tended to manage non-normally distributed data better, whilst earlier and smaller studies often did not, which makes the robustness of their findings questionable. The issue of multiple comparisons is solved in different ways: sometimes by using Bonferroni-type corrections (although most studies that use this method only control the false positive rate within a particular test, rather than across all the tests reported in the study), and sometimes by simply acknowledging the problem and reporting uncorrected statistics nevertheless. However, with the current body of literature meaning that many findings have been replicated, it is possible to feel more confident about which findings are consistent and which are less so. In addition, where studies make their mean (SD) results available, it is possible to compare effect sizes between studies.

4.1.5 Summary

Overall, the quality of the studies included in the review was variable, ranging from very small (<10 per group) studies, in which possible confounds were not always taken into account and data distribution was not considered in the analysis, to much larger (>100 per group) studies which were adequately powered to detect small effects, and adjusted for a number of potential confounds in their analysis. In terms of outcome, the most robust results are likely to be those from the studies that were adequately powered and in which consideration was given to data distribution and confounding variables prior to analysis. The use of a gene-negative control group is also likely to influence outcome. Based on this assessment, the most robust studies are likely to be the two large longitudinal studies, PREDICT-HD and Track-HD (Johnson et al., 2007; Tabrizi et al., 2009), as well as by those medium- or small-sized studies whose analysis took into account confounds, and data distribution

(Henley et al., 2008; Kipps et al., 2005; Snowden et al., 2008; Sprengelmeyer et al., 2006). In the discussion that follows, more weight is given to findings from those studies that were assessed as being of higher quality.

4.2 CONCLUSIONS

4.2.1 Facial emotion recognition

Using facial stimuli, anger recognition was most consistently impaired in manifest HD populations, closely followed by disgust and fear. Sadness and surprise were less consistently affected, and happiness very rarely. In premanifest populations there was sometimes no detectable deficit, sometimes facial disgust recognition was the only impairment seen, and in two studies impairment was seen across all negative emotions. Undoubtedly facial emotion recognition performance gets worse as populations move from being premanifest to manifest. Because either CAG repeat data or some measure of disease severity (e.g. UHDRS motor score, or disease duration) were not reported by a number of studies it is hard to ascertain whether an increasingly broad, or simply a more severe impairment is associated with more advanced disease (although some studies examine this within a single cohort (e.g., Johnson et al., 2007)). Three studies lend evidence to the suggestion that facial disgust recognition is the earliest emotion and modality to be affected, and that therefore deficits become broader as disease progresses (Gray et al., 1997; Hennenlotter et al., 2004; Sprengelmeyer et al., 2006). However the two studies judged to be of the highest quality report that wider deficits are detectable even in very early premanifest participants (Johnson et al., 2007; Tabrizi et al., 2009); one specifically states that disgust recognition is no more impaired than the other negative emotions (Johnson et al., 2007). This lends support to the idea that all

negative emotions are affected from an early stage, and that impairments then worsen across emotions as disease progresses. Certainly in manifest cohorts the majority of studies find no evidence that facial disgust recognition is disproportionately impaired compared with other negative emotions. Overall evidence suggests that negative emotions are more impaired than positive / ambiguous (happiness and surprise), and that this impairment progresses with disease, but can be detected in all negative emotions even in very early premanifest participants.

4.2.2 Vocal emotion recognition

Vocal emotion recognition has been studied far less, but the overall picture is similar. In manifest cohorts disgust recognition is consistently impaired, although anger and fear are often impaired as well. Two studies find that the disgust impairment is greater than that for other emotions (Hayes et al., 2007; Sprengelmeyer et al., 1996) but two find that anger or fear are affected more than disgust (Calder et al., 2010; Snowden et al., 2008). The one study investigating this in a premanifest group found no evidence of any deficits across emotions; thus as with faces there is no convincing evidence that vocal disgust recognition is affected more, or earlier, than other emotions. Unlike facial emotion recognition, no impairments have been found for vocal sadness recognition. On average, participants who could recognise vocal sadness were impaired at recognising facial sadness. It may be that sadness is easier to recognise vocally than facially, or that different modalities are differentially affected in HD. However, no study tested inter-modality differences statistically, and evidence of a significant group difference in one modality, but not in another, is not in itself evidence that performance in one modality is significantly different to performance in another. Further inter-modality work will be needed to clarify this.

4.2.3 Recognition of emotion in other modalities

Few studies looked at modalities other than faces or voices. Snowden et al. (2008) included an "eyes only" stimulus set, which in this review has been discussed separately from faces, because of the different information available in such stimuli. Whereas when using entire faces, disgust and anger recognition were impaired, when using only eyes anger recognition was unimpaired and sadness was impaired. This might suggest that angry faces are more easily disambiguated when non-eye information is hidden, whereas the opposite is true for sadness, which perhaps relies more on the lower half of the face. As this is a single study it would benefit from being replicated.

One study looked at body language and found no evidence of impaired emotion recognition (Aviezer et al., 2009). However some of the body language stimuli contained semantic clues (dirty underwear for disgust, and a gravestone for sadness). It is therefore possible that people with HD were able to label the images based on previously acquired semantic knowledge, rather than recognition of the emotion conveyed by the body postures of the models. One study looked at the odour and taste domains, although necessarily only tested disgust in these domains (Mitchell et al., 2005). After failing olfactory screening, two participants were dropped from the odour tests, but retained in the taste test, although there is evidence that flavour discrimination also depends on odours and therefore it is questionable whether these two participants should have remained in the study at all (Schiffman & Gatlin, 1993). This study shows that participants with HD tended to rate disgusting odours and taste combinations as less disgusting than controls. However this study differs from those in the facial and vocal domains to some extent as it is asking for subjective ratings, rather than categorisation of an external stimulus. Mean ratings suggest that

participants with HD were still able to discriminate between pleasant and less pleasant stimuli, so these results may represent a dulling of subjective sensations, rather than an inability to recognise objectively that odours and tastes might be categorised as "disgusting".

4.3 FUTURE RESEARCH

Future research would benefit from focusing on a number of areas. Firstly, improving study quality and consistency, for example by reporting sufficient demographic data to allow cohorts to be compared, and by careful consideration of possible confounds, and appropriate statistical analysis. There could also be more consistency with stimulus presentation (particularly in the facial recognition domain) although subtle differences in stimulus sets and presentation did not seem to lead to large differences in outcome in the studies reviewed here. Perhaps more important is the finding that using completely different (non-Ekman) stimulus sets results in different findings. Only one study attempted to use more ecologically valid facial stimuli (colour videos) and this seems to be a major gap in the field. In addition, most studies use verbal labelling as their response of choice (some studies use within-modality stimulus matching but this was not considered a sufficiently robust test of *emotion* recognition for inclusion in this review). It would be interesting to test between-modality stimulus-matching as an alternative to verbal labelling. Also, future work should certainly include tasks in more than one modality in order to test statistically whether impairments in one modality are different to those seen in another, as currently there is no clear evidence on this.

4.4 CLINICAL IMPLICATIONS

There is currently a large body of research investigating potential biomarkers and endpoints for clinical trials in HD. One of the implications from these findings is that emotion recognition may be of potential use as a biomarker. Emotion scores (both composites and individual emotions) are already included in the two largest ongoing longitudinal studies, PREDICT-HD (Paulsen et al., 2006) and Track-HD (Tabrizi et al., 2009) and further work should demonstrate how sensitive emotion recognition is at tracking decline over time, relative to other potential markers.

On an individual level, these results clearly show that in people with manifest HD emotion recognition tends to be impaired across negative emotions, in at least the facial and vocal domains, and that subtle impairments can be detected many years before motor onset. Anecdotally, carers and spouses often report that their partner with HD seems to react inappropriately when they are portraying emotions: one spouse told the first author that she was discussing something very upsetting with her husband, and he responded by talking about a new car he wanted to buy; another reported that whilst his wife got angry easily, she no longer seemed to recognise when he was angry, and to get her to appreciate his mood he needed to be explicit about his feelings and explain why he felt them. Whilst it seems clear that one of the main difficulties caused by poor emotion recognition would be social interaction, there is little formal evidence of how these difficulties impact on the relationships of people with HD. The research reviewed here also raises the question of whether the difficulties people with HD have with facial emotion recognition are as severe, or affect the same emotions, as those seen when non-verbal vocal stimuli, or prosodic stimuli are used. Both these areas would benefit from more research, firstly in order to outline the impact that these difficulties have, and secondly to investigate what

strategies might go some way towards overcoming them. For example, if people with HD can recognise an emotion better in one modality than another, carers could use this knowledge to help them judge how best to get their feelings across. Currently the literature does not address either of these questions adequately.

4.5 LIMITATIONS

This review was limited to peer-reviewed publications, meaning that grey literature and unpublished data were not included. Whilst this ensures some level of methodological rigour, in that all the studies have gone through the peer-review process, publication bias may mean that studies that did not find deficits in emotion recognition have been overlooked. Although efforts were made to clarify queries with corresponding authors of all the included studies, some were not contactable. This means that some data queries have gone unanswered, as have requests for extra demographic or outcome data.

5 CONCLUSION

The literature currently supports the conclusion that vocal and facial negative emotion recognition is impaired in both premanifest and manifest Huntington's disease, and suggests that impairments worsen with disease progression. There is yet to be convincing evidence of disproportionate impairments in particular emotions, and it is unclear whether impairments in one modality are of greater severity than those in another. Future work could usefully focus on developing more ecologically-valid stimulus sets and comparing performance between modalities directly, in order to answer some of the questions that still remain with regards to the deficits seen in this population. As neuropsychological metrics of clear relevance to patients' everyday lives, there is an overarching need further to evaluate emotion processing

measures as potential biomarkers for symptomatic and disease-modifying therapies in this devastating disease.

6 REFERENCES

Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr.Opin.Neurobiol., 12,* 169-177.

Aviezer, H., Bentin, S., Hassin, R. R., Meschino, W. S., Kennedy, J., Grewal, S. et al. (2009). Not on the face alone: perception of contextualized face expressions in Huntington's disease. *Brain, 132,* 1633-1644.

Benton, A. L., Sivan, A. B., Hamsher, K. d., Varney, N. R., & Spreen, O. (1978). *Benton Facial Recognition Test*. Lutz, FL: Psychological Assessment Resources Inc.

Calder, A. J., Keane, J., Lawrence, A. D., & Manes, F. (2004). Impaired recognition of anger following damage to the ventral striatum. *Brain, 127,* 1958-1969.

Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I. et al. (2003). Facial expression recognition across the adult life span. *Neuropsychologia, 41,* 195-202.

Calder, A. J., Keane, J., Young, A. W., Lawrence, A. D., Mason, S., & Barker, R. A. (2010). The relation between anger and different forms of disgust: implications for emotion recognition impairments in Huntington's disease. *Neuropsychologia, 48,* 2719-2729.

Calder, A. J., Young, A. W., Rowland, D., Perrett, D. I., Hodges, J. R., & Etcoff, N. L. (1996). Facial emotion recognition after bilateral amygdala damage: Differentially severe impairment of fear. *Cognitive Neuropsychology, 13,* 699-745.

Ekman, P. & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P. (1992). Are there basic emotions? *Psychol Rev., 99,* 550-553.

Ekman, P. (1993). Facial expression and emotion. *Am Psychol, 48,* 384-392.

Gao, X. & Maurer, D. (2010). A happy story: Developmental changes in children's sensitivity to facial expressions of varying intensities. *J.Exp.Child Psychol., 107,* 67-86.

Gray, J. M., Young, A. W., Barker, W. A., Curtis, A., & Gibson, D. (1997). Impaired recognition of disgust in Huntington's disease gene carriers. *Brain, 120,* 2029-2038.

Harper, P. (2002). The epidemiology of Huntington's disease. In G.Bates, P. Harper, & L. Jones (Eds.), *Huntington's Disease* (3 ed., pp. 3-27). Oxford: Oxford University Press.

Hayes, C. J., Stevenson, R. J., & Coltheart, M. (2007). Disgust and Huntington's disease. *Neuropsychologia, 45,* 1135-1151.

Hayes, C. J., Stevenson, R. J., & Coltheart, M. (2009). The processing of emotion in patients with Huntington's disease: variability and differential deficits in disgust. *Cogn Behav.Neurol., 22,* 249-257.

Henley, S. M., Bates, G. P., & Tabrizi, S. J. (2005). Biomarkers for neurodegenerative diseases. *Curr.Opin.Neurol, 18,* 698-705.

Henley, S. M., Wild, E. J., Hobbs, N. Z., Warren, J. D., Frost, C., Scahill, R. I. et al. (2008). Defective emotion recognition in early HD is neuropsychologically and anatomically generic. *Neuropsychologia, 46,* 2152-2160.

Hennenlotter, A., Schroeder, U., Erhard, P., Haslinger, B., Stahl, R., Weindl, A. et al. (2004). Neural correlates associated with impaired disgust processing in pre-symptomatic Huntington's disease. *Brain, 127,* 1446-1453.

Huntington Study Group (1996). Unified Huntington's Disease Rating Scale: reliability and consistency. Huntington Study Group. *Mov Disord, 11,* 136-142.

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.

Johnson, S. A., Stout, J. C., Solomon, A. C., Langbehn, D. R., Aylward, E. H., Cruce, C. B. et al. (2007). Beyond disgust: impaired recognition of negative emotions prior to diagnosis in Huntington's disease. *Brain, 130,* 1732-1744.

Kipps, C. M., Duggins, A. J., Mahant, N., Gomes, L., Ashburner, J., & McCusker, E. A. (2005). Progression of structural neuropathology in preclinical Huntington's disease: a tensor based morphometry study. *J Neurol Neurosurg Psychiatry, 76,* 650-655.

Kipps, C. M., Duggins, A. J., McCusker, E. A., & Calder, A. J. (2007). Disgust and happiness recognition correlate with anteroventral insula and amygdala

volume respectively in preclinical Huntington's disease. *J Cogn Neurosci, 19,* 1206-1217.

Kremer, B. (2002). Clinical neurology of Huntington's disease. In G.Bates, P. Harper, & L. Jones (Eds.), *Huntington's Disease* (3 ed., pp. 28-61). Oxford: Oxford University Press.

Langbehn, D. R., Brinkman, R. R., Falush, D., Paulsen, J. S., & Hayden, M. R. (2004). A new model for prediction of the age of onset and penetrance for Huntington's disease based on CAG length. *Clin Genet., 65,* 267-277.

Mahant, N., McCusker, E. A., Byth, K., & Graham, S. (2003). Huntington's disease: clinical correlates of disability and progression. *Neurology, 61,* 1085-1092.

Milders, M., Crawford, J. R., Lamb, A., & Simpson, S. A. (2003). Differential deficits in expression recognition in gene-carriers and patients with Huntington's disease. *Neuropsychologia, 41,* 1484-1492.

Mitchell, I. J., Heims, H., Neville, E. A., & Rickards, H. (2005). Huntington's disease patients show impaired perception of disgust in the gustatory and olfactory modalities. *J Neuropsychiatry Clin Neurosci, 17,* 119-121.

Montagne, B., Kessels, R. P., Kammers, M. P., Kingma, E., de Haan, E. H., Roos, R. A. et al. (2006). Perception of Emotional Facial Expressions at Different Intensities in Early-Symptomatic Huntington's Disease. *Eur Neurol, 55,* 151-154.

Novak, M. J. & Tabrizi, S. J. (2010). Huntington's disease. *BMJ, 340,* c3109.

Paulsen, J. S., Hayden, M., Stout, J. C., Langbehn, D. R., Aylward, E., Ross, C. A. et al. (2006). Preparing for preventive clinical trials: the Predict-HD study. *Archives of Neurology, 63,* 883-890.

Paulsen, J. S., Langbehn, D. R., Stout, J. C., Aylward, E., Ross, C. A., Nance, M. et al. (2008). Detection of Huntington's disease decades before diagnosis: The Predict HD study. *J.Neurol.Neurosurg.Psychiatry, 79*, 874-880.

Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V. et al. (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proc.Biol.Sci., 265,* 1809-1817.

Rosenblatt, A., Liang, K. Y., Zhou, H., Abbott, M. H., Gourley, L. M., Margolis, R. L. et al. (2006). The association of CAG repeat length with clinical progression in Huntington disease. *Neurology, 66,* 1016-1020.

Rubinsztein, D. C., Leggo, J., Coles, R., Almqvist, E., Biancalana, V., Cassiman, J. J. et al. (1996). Phenotypic characterization of individuals with 30-40 CAG repeats in the Huntington disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36-39 repeats. *Am J Hum.Genet., 59,* 16-22.

Schiffman, S. S. & Gatlin, C. A. (1993). Clinical physiology of taste and smell. *Annu.Rev.Nutr., 13,* 405-436.

Scott, S. K., Young, A. W., Calder, A. J., Hellawell, D. J., Aggleton, J. P., & Johnson, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature, 385,* 254-257.

Snowden, J. S., Austin, N. A., Sembi, S., Thompson, J. C., Craufurd, D., & Neary, D. (2008). Emotion recognition in Huntington's disease and frontotemporal dementia. *Neuropsychologia, 46,* 2638-2649.

Sprengelmeyer, R., Schroeder, U., Young, A. W., & Epplen, J. T. (2006). Disgust in pre-clinical Huntington's disease: A longitudinal study. *Neuropsychologia, 44,* 518-533.

Sprengelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., Homberg, V. et al. (1996). Loss of disgust - perception of faces and emotions in Huntington's disease. *Brain, 119,* 1647-1665.

Stout, J. C., Paulsen, J. S., Queller, S., Solomon, A. C., Whitlock, K. B., Campbell, J. C. et al. (2011). Neurocognitive signs in prodromal Huntington disease. *Neuropsychology, 25,* 1-14.

Tabrizi, S. J., Langbehn, D. R., Leavitt, B. R., Roos, R. A., Durr, A., Craufurd, D. et al. (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.

Young, A. W., Perrett, D., Calder, A., Sprengelmeyer, R., & Ekman, P. (2002). *Facial expressions of emotion: stimuli and tests (FEEST).* Bury St. Edmunds: Thames Valley Test Company, Suffolk.

PART 2: EMPIRICAL PAPER

EMOTION RECOGNITION IS IMPAIRED ACROSS MODALITIES IN MANIFEST HUNTINGTON'S DISEASE

ABSTRACT

Aims

Huntington's disease (HD) is known to cause impaired emotion recognition, particularly in the facial domain. Emotion recognition in other domains has been less well studied. No study has compared formally emotion recognition performance in more than one modality. The aim of this study was to assess emotion recognition in early HD using a range of modalities (facial, vocal and musical), in order to examine whether impairments were cross-modal.

Methods

Twenty-five participants with early HD (CAG repeat length>38) were compared with 25 neurologically-normal controls on a range of measures including estimated pre-morbid IQ, executive function, facial recognition, music perception, and emotion recognition. In each of three modalities (faces, voices, music) 10 stimuli for each of four emotions were presented and participants asked to select a response from "happy", "sad", "angry" or "fearful".

Results

After adjusting for age, estimated pre-morbid IQ, facial recognition skills, executive function and gender, the HD group was impaired relative to controls at recognising sad, angry and fearful stimuli in all three modalities. There was a tendency for the HD group to find fearful faces and voices harder to recognise than sad or angry faces and voices, after adjusting for control performance. There was also evidence that the HD group found fearful faces and voices harder to recognise than fearful music.

Conclusions

The emotion recognition impairment in HD appears to be cross-modal, suggesting the involvement of relatively high-level neural systems, rather than modality-specific mechanisms. Possible neural substrates are discussed, as well as the implications for the symptomatic management of the disease.

1 INTRODUCTION

Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder caused by an expanded CAG repeat on chromosome 4. It is classically characterised by involuntary movements, and cognitive and psychiatric deficits, with the onset of motor signs usually occurring in mid-adulthood. Although for many years it was characterised primarily as a movement disorder, it is now well recognised that cognitive and psychiatric problems accompany and frequently precede motor signs (Lawrence et al., 1998; Snowden, Craufurd, Thompson, & Neary, 2002). It has a prevalence in the UK of ~1 per 10,000 individuals with many more at risk of the disease (Harper, 2002). It progresses slowly and inexorably, with increasingly distressing and disabling symptoms associated with considerable care-giver and family burden, with morbidity and death occurring 15-20 years from onset, by which time the patient is often bed-bound and mute.

Recognising emotions in others is an important social skill, and much work has been done on studying the expression and recognition of canonical emotions which appear to be cross-cultural, and which it is argued have a biological basis (Ekman, 1992). Deficits in emotion recognition have been reported in both premanifest (prior to motor onset, PM) and early HD. Originally facial disgust recognition was thought to be particularly impaired (see, for example, Gray, Young, Barker, Curtis, & Gibson, 1997; Montagne et al., 2006; Sprengelmeyer et al., 1996). It has also been suggested that disgust recognition is similarly impaired in the vocal, olfactory and gustatory domains (Hayes, Stevenson, & Coltheart, 2007; Mitchell, Heims, Neville, & Rickards, 2005; Sprengelmeyer et al., 1996). However more recently larger studies have failed to replicate the finding of disproportionately impaired facial disgust recognition, showing instead an impairment across negative emotions (sadness,

disgust, anger and fear) and occasionally showing most impairment at fear or anger recognition (Aviezer et al., 2009; Henley et al., 2008; Johnson et al., 2007; Milders, Crawford, Lamb, & Simpson, 2003; Snowden et al., 2008).

A systematic review of the literature (Henley et al., in press) supports the conclusion that in manifest HD emotion recognition is impaired fairly consistently across all negative emotions in the facial domain. Far fewer studies have investigated vocal emotion recognition, and those that have find that disgust recognition is impaired most often (four out of four studies) but that anger and fear recognition are often impaired as well. Although some studies have looked at facial and vocal emotion recognition in the same cohort, to date no study has carried out direct statistical comparisons of performance in different modalities, so it is not clear whether cross-modal deficits are of similar magnitudes. In addition, very few studies have attempted to assess emotion recognition in modalities other than facial and vocal.

It is unclear why relatively few studies have focused on vocal emotion recognition in HD, nor why there have been no direct comparisons of facial and vocal emotion recognition. Deficits in recognising emotional prosody in HD were first reported in 1990 (Speedie, Brake, Folstein, Bowers, & Heilman) and one of the earliest studies investigating emotion-specific impairments in HD included both facial and vocal stimuli (Sprengelmeyer et al., 1996). As with facial expressions, canonical vocal emotional expressions are recognised cross-culturally, and very rapidly (Sauter & Eimer, 2010; Sauter, Eisner, Ekman, & Scott, 2010). Vocal and facial emotion recognition of specific emotions are thought to depend on similar neural substrates including, for example, insula-striatal systems for disgust, amygdala for fear, ventral striatum for anger (Adolphs, 2002; Calder, Keane, Lawrence, & Manes, 2004;

Calder, Keane, Manes, Antoun, & Young, 2000; Calder, Lawrence, & Young, 2001). Within HD research several studies have used structural or functional imaging to examine the neural substrates underlying facial emotion recognition (Henley et al., 2008; Hennenlotter et al., 2004; Kipps, Duggins, McCusker, & Calder, 2007) and have found similar brain regions to be implicated. It would therefore be predicted that the facial emotion recognition deficits seen in HD should also be seen in other modalities (including vocal); indeed based on current neurological evidence it might seem unusual if such deficits were not cross-modal.

Music is another modality through which emotion can be expressed (Krumhansl, 1997). Musical stimuli have tended to be underrepresented in formal studies of emotion recognition, perhaps because of a belief that the ability to perceive emotions in music is a relatively recent phenomenon in human development, based on purely aesthetic features, rather than evolving as a useful survival mechanism as recognition of emotions in faces and voices may have done. However, there is increasing evidence that listening to music is associated with activation patterns in brain regions that are implicated in general reward and emotion perception, such as the ventral striatum, amygdala, prefrontal cortex and insula (Blood & Zatorre, 2001; Brown, Martinez, & Parsons, 2004; Griffiths, Warren, Dean, & Howard, 2004). In a recent review Koelsch (2010) argues that in fact music emotion recognition has evolved to have social "survival" value, for example helping humans to achieve goals of communication, collaboration, and playing a role in social cohesion and reciprocal care; hence it is unsurprising that recognition of emotion in a relatively abstract stimulus (music) seems to use, in a large part, similar neural networks as recognition of emotions in other modalities. Indeed, several of these brain areas (the so called "social brain") are consistently activated in theory of mind and empathy tasks, tasks

in which participants are required to represent the state of the world around them, and emulate the feelings of others (Dodell-Feder, Koster-Hale, Bedny, & Saxe, 2011; Singer, 2006). Music is also by its nature a relatively abstract representation of emotion and may therefore depend on more abstract reasoning abilities (and underlying neural substrates) than face and vocal emotion processing.

People with little or no formal musical training are consistent, and very rapid, in their emotional ratings of classical music into basic emotional categories (Peretz, Gagnon, & Bouchard, 1998). There is also some evidence that even when musical perceptual processing is impaired by acquired brain damage (e.g. the ability to recognise excerpts, and detect differences in pitch and rhythm are lost), musical affective processing can be spared (Peretz et al., 1998). This suggests that it is not necessary to be either highly trained, or to have intact musical perceptual systems in order to appreciate the emotional content of music. In addition, there is at least one report of an individual whose musical perceptual processing was within normal limits, but who reported an inability to experience as pleasurable music he had enjoyed prior to a left hemisphere infarct (Griffiths et al., 2004). This lends support to the idea that systems for processing the affective value of music might be functionally and anatomically distinct from those needed to process the perceptual characteristics.

The fact that similar neural substrates may underlie both music emotion processing and recognition of emotions in more biologically-relevant stimuli such as faces and voices suggests that music emotion recognition may merit further examination. In populations with a neurodegenerative disease, many of the brain regions implicated in emotion recognition are compromised, and deficits in music emotion recognition have previously been reported in populations with fronto-temporal lobar

degeneration (Omar et al., 2011) and Parkinson's disease (van Tricht, Smeding, Speelman, & Schmand, 2010).

Given the locus of atrophy in early HD (marked in the striatum, with evidence that frontal and other cortical areas are more mildly affected until later stages (see e.g., Tabrizi et al., 2009)), and previous behavioural evidence, we would predict that emotion recognition would be impaired across facial, vocal and musical modalities. However, as mentioned above, the extent to which impairments in one modality are comparable to those in another has not yet been ascertained. In HD, the question of the extent to which any deficit seen is uniform across modalities is an important one. Not only does this have implications for the neural substrates involved, but would also impact on potential symptomatic treatments. Although there is little formal research in this area, anecdotally carers often report how distressing they find it when their emotions go unrecognised and unacknowledged by people with HD. If recognition of some emotions was shown to be spared in one modality relative to another, information from the former modality might be used successfully to partially alleviate the difficulties people with HD have in the others. The aim of the current study was therefore to compare directly facial, vocal and musical emotion recognition in the same cohort of HD participants. This could have important implications for our understanding of the experience of people with HD, and for symptomatic treatment of the disease, as well as adding to the body of evidence about the similarities between emotion recognition performance with faces and voices, and the more abstract stimulus of music.

2 METHODS

2.1 PARTICIPANTS

Twenty-five patients with genetically-confirmed HD were recruited from the multidisciplinary HD clinic at the National Hospital for Neurology and Neurosurgery, London, the Institute of Neurology HD Research Database and the Huntington's Disease Clinic at Addenbrooke's Hospital, Cambridge. All patients had a CAG repeat length of >38 and were clinical stage 1 or 2 as determined by functional assessment (Shoulson & Fahn, 1979). Twenty-five neurologically-normal controls, who comprised either the patients' partners or at-risk subjects who had tested negative for the HD gene expansion, were also recruited. Such controls are more likely to experience similar social and environmental influences to HD gene carriers than healthy volunteers from the general population, and therefore these factors are less likely to play a role in group differences. Participants with a history of substance abuse, estimated pre-morbid IQ of below 80, or significant medical, neurological or psychiatric comorbidity (non-HD related) were excluded. No participant reported a hearing difficulty. All participants gave written informed consent in accordance with the Declaration of Helsinki, and the study had local research ethics committee and UCLH NHS Trust approval (see Appendix 3).

Twenty-five participants (16 early HD and nine controls) were initially recruited as part of the London Longitudinal HD study and the relevant neuropsychological data were collected, by the author, as part of that study (see e.g., Henley et al., 2009). However the emotion recognition data presented here have not been presented elsewhere prior to this. An additional 25 participants were recruited for the current study.

2.2 SAMPLE SIZE

Previous studies of facial emotion recognition have found large (d>0.8) effect sizes (Cohen, 1992) when comparing recognition of fear and anger between early HD and controls using the Ekman pictures of facial affect (e.g., Henley et al., 2008; Milders et al., 2003), when using vocal stimuli (Sprengelmeyer et al., 1996) (although this study used a different set of stimuli to those in the current study) and when using the music emotion stimuli described here (Omar et al., 2011) (in a cohort with fronto-temporal lobar degeneration). Most studies fail to find evidence of an impairment in happiness or sadness recognition (in any modality) although it was deemed important to include these stimuli as comparisons and to keep the nature of the test similar to previous work.

For the purposes of this study sample size calculations focused on anger and fear recognition in the three modalities and were based on effect sizes derived from the above studies. These studies differed in key ways from the one proposed here (particularly in the nature of the behavioural test, and the demographics of the cohorts tested) and this must be taken into account when estimating sample sizes.

Sample sizes were calculated for logistic regression with a single binary predictor (in this case group) with α=0.05 (one-tailed) and power=0.7 using GPower 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007). Given the exploratory nature of the study and the fact that this test is novel, 70% power was deemed appropriate at this stage. If effect sizes turn out to be as predicted then follow-up work could be planned with increased power. The correlation between gender, age, IQ and group in the above studies was minimal (and should be in the planned study) and so the variance inflation factor (R^2 for other covariates) was estimated at 0.05 (Hsieh, Bloch, &

Larsen, 1998). Based on the above studies, estimated proportions correct were 90% in controls, and 50% in HD for facial and vocal stimuli, and 75% in controls and 30% in HD for music stimuli. Estimates for sample sizes in each group were 24 (music stimuli) and 25 (facial and vocal stimuli) (Faul et al., 2007). Given the exploratory nature of the study, a sample size of 25 in each group was therefore decided on. This should mean that the study is powered adequately to detect statistically significant group differences in emotion recognition, assuming large effect sizes.

2.3 MEASURES

All HD participants had previously been assessed by neurologists using the standard neurological test for HD (Unified Huntington's disease rating scale, UHDRS (Huntington Study Group, 1996)) and consented to these data being used in the current study to provide information about their disease state. The current battery included tests of facial, vocal, and musical emotion recognition, as well as background tests of potential confounding variables such as estimated pre-morbid IQ, executive function, face recognition, music perception and musical knowledge as follows:

2.3.1 Estimated-premorbid IQ

Pre-morbid IQ was estimated using the National Adult Reading Test (NART) (Nelson & Willison, 1991).

2.3.2 Executive function

The Trail-Making Test (Reitan & Wolfson, 2004) (time to complete Part B minus time to complete Part A) was used as a measure of "executive function", in this case set-switching, which is known to be affected in HD.

2.3.3 Facial recognition

Face recognition was tested using the short form of the Benton Facial Recognition Test (Benton, Hamsher, Varney, & Spreen, 1983).

2.3.4 Musical perception

Music perception was assessed using the Montreal Battery of Evaluation of Amusia (MBEA), for which age-matched normative data are available for musically untrained subjects (Peretz, Champod, & Hyde, 2003). This battery is based on a two alternative (same/different) forced choice comparison of pairs of short unfamiliar musical sequences. Four subtests of the MBEA were used: scale (key), pitch contour (melody), pitch interval, and rhythm. As this battery was not available to participants who were assessed as part of the London Longitudinal HD study, it was only possible to assess music perception in a subset of participants (14 controls and eight early HD) who were recruited more recently.

2.3.5 Musical background

Musical background was measured using the Hailstone Questionnaire (Hailstone et al., 2009) (see also Appendix 4). This questionnaire was not available for five of the earliest-tested London Study participants. Data on participants' musical background is presented in Appendix 5.

2.3.6 Emotion recognition

A novel battery was designed to assess recognition of four emotions (happiness, sadness, anger, fear) in music, for comparison with recognition of these emotions from facial expression and nonverbal vocal sounds. The target emotions chosen represent four of the six canonical emotions in the original set of emotional faces created by Ekman and Friesen (1976); surprise and disgust were excluded due to the difficulty of creating musical equivalents for these. The novel battery has previously been described and used with participants with a diagnosis of fronto-temporal lobar degeneration (Omar et al., 2011; Omar, Hailstone, Warren, Crutch, & Warren, 2010).

2.3.6.1 Stimuli: music

The stimuli for recognition of emotion in music were excerpts drawn from the Western classical canon and film scores (mean duration (range) as follows: anger 11.6 sec (9.8 – 13.3); fear, 12.2 sec (10.3 – 16.4); happiness, 10.5 sec (8 – 13.3); sadness, 11.6 sec (10.1 – 16.0)). Stimuli were selected for inclusion in the battery based on an initial pilot study (described in Appendix 6) in 16 healthy participants who did not participate in any subsequent experiments. Most pieces were orchestral works; some chamber pieces are also included. Songs and other vocal musical pieces were excluded to avoid confounding from primarily vocal emotion processing. Stimuli are available on the CD submitted with this thesis.

2.3.6.2 Stimuli: facial expressions

The facial emotion stimuli comprised black and white photographs of posed facial expressions derived from the set produced by Ekman and Friesen (1976); the most reliably recognised exemplars from the original set for each target emotion were selected.

2.3.6.3 Stimuli: nonverbal vocal sounds

The vocal emotion stimuli were brief nonverbal vocalisations recorded by male and female actors to express each of the same target canonical emotions (Sauter, 2006). The most reliably recognised exemplars from the original set for each target emotion were selected.

2.3.6.4 General testing procedure

Stimulus presentation was on a notebook computer in a quiet room, free from distraction. Cogent 2000 (www.vislab.ucl.ac.uk/Cogent2000) running under MATLAB 7.0® (http://www.mathworks.com) was used for stimulus presentation, with participant responses collected for off-line analysis.

Each trial consisted of stimulus presentation, with simultaneous presentation of the four target emotion words in a random order at the corners of the screen, i.e. a four-alternative forced-choice (4AFC) paradigm. Visual stimuli (Ekman faces) were centred on the screen, and remained on screen until a response was made. Auditory stimuli were presented as digital wave files in free field at a comfortable listening level. Participants could choose to repeat auditory stimuli as many times as they wished prior to making a response. In both visual and auditory trials the target emotion words remained on screen throughout the trial. Vocal stimuli were on average two seconds long, and musical stimuli were on average 11 seconds long. Each target emotion word was presented next to a number from 1-4. Participants responded orally when they had made a choice, and the examiner entered the numbered button-press corresponding to the participant's response. Examiner button-press, rather than participant button-press, was used because prior experience with this population showed that the motor symptoms of HD could in cases affect

button-press accuracy. A trial ended only when the participant choice had been entered by the examiner (i.e. trials were not time-limited). At this point a blank screen was displayed for 2 seconds prior to the next trial.

For each modality 40 trials were presented, comprising 10 stimuli representing each of the four target canonical emotions. Modalities were presented in a block design, in the order: faces, vocal sounds, music. Within each modality (block), the 40 trials were presented in randomised order. For the music block participants were also asked whether the stimulus was familiar or not. Before the start of each modality block, four practice trials were administered to ensure the participant understood the task. No feedback about performance was given during the test.

2.4 PROCEDURES

Neuropsychological tasks were administered by the same investigator in a single session lasting approximately 2 ½ hours, either at UCL Institute of Neurology, or at the participant's home.

2.5 STATISTICAL ANALYSIS

Data were analysed using STATA version 9.2 (Stata Corporation, College Station, Texas, 2006).

2.5.1 Demographic and background data

A Chi-square test was used to examine whether gender differed between groups and Fisher's exact test was used to examine whether handedness differed between groups. t-tests, allowing for unequal variance where necessary, were used to investigate group differences in age and estimated premorbid IQ.

Fisher's exact test was used to compare the proportion in each group that had failed the Benton test and the Peretz test. A t-test, allowing for unequal variance, was used to compare group means on the Benton (this test is already age- and education-adjusted).

As the Trail-Making scores were not Normally distributed, linear regression models were used to assess group differences, with 95% accelerated bias-corrected bootstrap confidence intervals, with 2000 replicates, adjusting for the effects of age and estimated pre-morbid IQ by including them as covariates. A linear regression model was used to assess group differences in Peretz score, again adjusting for the effects of age and estimated pre-morbid IQ by including them as covariates. Adjusted between-group differences are reported here, with unadjusted between-group differences reported in Appendix 7 for comparison. The model assumes that adjusted between-group differences are constant for all levels of the covariates adjusted for, i.e. the adjusted between-group differences reported are independent of the level of the nuisance covariate.

2.5.2 Emotion recognition data

2.5.2.1 Differences between groups

Generalised linear mixed regression models were used to investigate the influence of group, modality, target emotion and their interactions on the probability of a correct response. A logistic link was used as is standard for binary outcomes (intended versus other response) and subject identity was included as a random effect to allow for associations between responses from the same individuals. The effects of age, estimated pre-morbid IQ, facial recognition ability (Benton), executive skills (Trail-Making Tests B minus A) and gender were adjusted for by including these variables

as covariates in the model. Adjusted odds ratios are reported here, with unadjusted odds ratios reported in Appendix 7. Within each modality (faces, voices and music) and emotion (happy, sad, angry, fearful) the odds of an intended response relative to that for controls were estimated (i.e. performance in the early HD group was compared with that of controls, for each emotion within each modality).

2.5.2.2 Modality- and emotion-dependent differences in the HD group

Where there was evidence of impairment in the early HD group relative to controls further planned comparisons were carried out. Recognition of different emotions was compared within each modality, to investigate whether there was evidence that some emotions were harder to recognise than others. Recognition of each emotion was also compared across the three modalities, to investigate whether there was evidence that recognition of a given emotion was harder in some modalities than others. For all these secondary analyses control performance was taken into account, to allow for normal differences in the recognition of different emotions in different modalities; thus, for example, the odds of an intended angry response to faces in early HD *relative to controls* was compared with the odds of an intended fear response to faces in the early HD group *relative to controls*.

Significance levels were not adjusted to take into account the number of comparisons because all the associations being investigated were thought to be of independent scientific interest (Rothman, 1990).

3 RESULTS

3.1 DEMOGRAPHIC DATA

Demographic data are shown in Table 3-1. Differences in gender, age, handedness and estimated pre-morbid IQ between the HD and control groups were small and non-significant (all $p>0.05$). The majority (23) of HD patients underwent neurological assessment within the year prior to cognitive testing. Neurological data for two HD patients were only available from 1.5 and 2 years prior to cognitive testing and are therefore likely to underestimate severity in these two patients.

Table 3-1 Demographic data

	Control (N=25)	HD (N=25)
Gender (M:F)	13:12	11:14
Age, years	48.9 (13.7)	54.0 (10.6)
Estimated premorbid IQ	109.8 (12.2)	108.7 (11.7)
Handedness (R:L)	23:2	21:4
CAG repeat length		42.8 (2.0), range 39 - 46
Duration of motor signs, years		6.4 (3.5)
UHDRS Motor		34.5 (14.9)
UHDRS Independence		88.6 (11.9)
UHDRS Total Functional Capacity		10.2 (2.2)

Data are shown as mean (s.d.) with the exception of gender and handedness; handedness was taken as the hand used to write with; UHDRS = Unified Huntington's Disease Rating Scale: motor is out of 124, higher score = more severely impaired; independence is scored as a percentage, higher score = better function; Total Functional Capacity is out of 13, higher score = better function

3.2 BACKGROUND DATA

Five HD participants failed the Benton Facial Recognition Test, compared with no controls, a result of borderline statistical significance (p=0.050, Fisher's exact test). On average the HD group scored significantly lower than controls on the Benton (mean difference 3.0 points, 95% CI -5.3, -0.7, $t(48)=2.60$, p=0.013). After adjusting for age and IQ the HD group was also slower, on average, to complete the Trail-Making Test A and Trail-Making Test B, than controls: TMTA mean difference 14.2 seconds, Bootstrap 95% CI 9.3, 21.4, $p<0.05$; TMT B mean difference 70.9 seconds, Bootstrap 95% CI 43.9, 108.8, $p<0.05$. After adjusting for age and IQ the HD group was also slower than controls at the difference between TMTA and B, mean difference 56.7 seconds, Bootstrap 95% CI 29.5, 87.4, $p<0.05$.

Two out of 14 controls, and three out of eight HD participants failed the Peretz test, which was not statistically significant (p=0.31, Fisher's exact test). However the two controls who failed scored just below the cut-off, whilst the three HD participants who failed scored very poorly; after adjusting for age and IQ the mean difference between groups was statistically significant: mean difference 12.8 points, 95% CI 3.4, 22.2, $t(18)=-2.85$, p=0.011.

Mean performance for these tasks, with unadjusted and adjusted group differences, are presented in more detail in Appendix 7. Further details of the regression models used are in Appendix 8.

3.3 EMOTION RECOGNITION

The mean proportions of correct responses for each target emotion are displayed in Figure 3.1. Adjustment was made by using the parameters from the model to predict the expected value of the outcome variable when all potential confounding variables were set to their mean levels in the entire cohort. Table 3-2 shows the odds ratios for a correct (intended) response for each emotion in each modality in the early HD group, relative to the odds of a correct (intended) response for the controls for each emotion in each modality, adjusted for age, estimated premorbid IQ, Benton score, Trail-Making B-A and gender. Unadjusted odds ratios are reported in Appendix 7 and further details of the regression model are in Appendix 8.

3.3.1 Faces

Early HD participants were significantly worse at recognising sad, angry and fearful faces relative to controls. Note that because of the ceiling effect observed for happy faces, happy faces were not included in the model.

3.3.2 Voices

Early HD participants were significantly worse at recognising sad, angry and fearful voices relative to controls, but not happy voices.

3.3.3 Music

Early HD participants were significantly worse at recognising sad, angry and fearful music relative to controls, but not happy music.

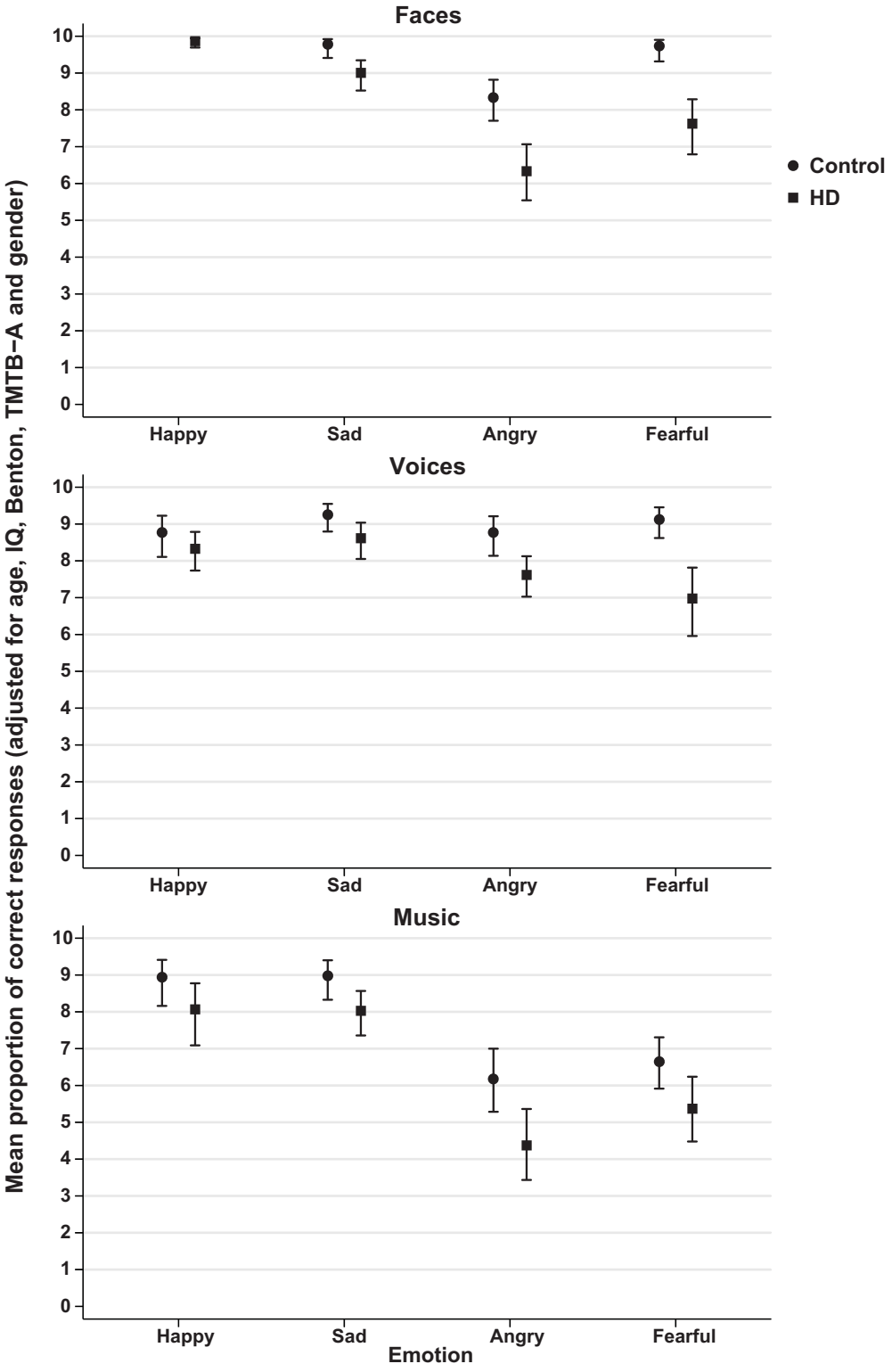

Figure 3.1 Mean proportion correct for each group, emotion and modality, adjusted for age, IQ, Benton, Trail-Making B-A score and gender, with 95% CIs

Table 3-2 Odds of a correct response (95% confidence intervals) relative to controls for each emotion in each modality, after adjusting for age, estimated pre-morbid IQ, Benton facial recognition score, Trail-Making B-A, and gender

	Intended emotion			
Faces	Happy	Sad	Angry	Fearful
Controls	[a]	1	1	1
Early HD	[a]	0.20 (0.07, 0.62)	0.35 (0.20, 0.60)	0.09 (0.03, 0.26)
		p=0.005	p<0.001	P<0.001

	Intended emotion			
Voices	Happy	Sad	Angry	Fearful
Controls	1	1	1	1
Early HD	0.70 (0.37, 1.34)	0.50 (0.26, 0.98)	0.45 (0.25, 0.81)	0.22 (0.11, 0.44)
	p=0.28	p=0.043	p=0.007	p<0.001

	Intended emotion			
Music	Happy	Sad	Angry	Fearful
Controls	1	1	1	1
Early HD	0.50 (0.21, 1.16)	0.46 (0.23, 0.93)	0.48 (0.28, 0.82)	0.58 (0.36, 0.95)
	p=0.10	p=0.031	p=0.007	p=0.028

An odds ratio <1 means that the early HD group is less likely to recognise stimuli correctly than controls; where 95% confidence intervals are both <1 this means that the odds ratio is statistically significant at the p<0.05 level.
a As controls were at ceiling on happy faces these data were excluded from the model, meaning there can be no comparison between groups for happy faces

3.3.4 Control performance

The modal response in controls across modalities and emotions was almost always that of the target emotion, i.e. in general most controls correctly identified the target emotion. Exceptions were one of the "angry" Ekman faces, which controls more often identified as "sad", and one of the "fearful" musical excerpts, which controls more often identified as "angry".

Controls were at ceiling for happy faces, and on average scored highly (>8/10 correct) for all other facial and vocal stimuli, across emotions. Happy and sad musical stimuli were equally well-identified by this group, but their performance tended to be lower for angry and fearful musical stimuli, implying that these emotions in this modality were inherently harder to recognise.

3.3.5 Within-group differences between emotions in each modality

After taking control levels of performance into account, there was evidence that early HD participants found fearful face recognition harder than angry face recognition (OR 0.25, 95% CI 0.08, 0.78, p=0.017); there were no other statistically significant differences between emotion recognition performance in this modality. In the vocal modality there was a tendency for HD participants to find fearful voices harder to recognise than angry voices (OR 0.49, 95% CI 0.24, 1.01, p=0.052). In the music modality there were no statistically significant differences between recognition of sad, angry or fearful stimuli in the HD group.

3.3.6 Within-group differences between modalities for each emotion

After taking control performance into account there was no evidence that recognition of sad or angry stimuli in the HD group differed depending on modality. On average, HD participants were worse at recognising fearful faces and fearful voices

than fearful music (fearful faces vs. music: OR 0.15, 95% CI 0.05, 0.44, p=0.001; fearful voices vs. music: OR 0.38, 95% CI 0.19, 0.74, p=0.004).

3.3.7 Repeating analysis without participants who failed Peretz test

It was not possible to include the Peretz score as a covariate in the main analysis as only a subset of participants completed the Peretz test and this would have left the analysis underpowered. However the emotion recognition analysis was repeated without the five participants who failed the Peretz test (two controls, three early HD). Without these participants there was no longer evidence that the HD group found sad voices or sad music harder to recognise than controls. Within the HD group there was now evidence that these participants found fearful voices harder to recognise than both sad and angry voices (fearful vs. sad voices: OR 0.32, 95% CI 0.12, 0.83, p=0.019; fearful vs. angry voices: OR 0.41, 95% CI 0.19, 0.89, p=0.024); in the main analysis there was merely a trend towards this effect in fearful vs. angry voices. Other findings were not materially changed.

4 DISCUSSION

4.1 MAIN FINDINGS

This study demonstrates that early HD patients have difficulty recognising sad, angry and fearful stimuli in the facial, vocal and musical modalities. Fearful stimuli tended to be harder for this group to recognise than sad or angry stimuli, in both the facial and vocal modalities. There was also evidence that fear was harder for the HD group to recognise when presented using facial or vocal stimuli, compared with musical stimuli, although there was no evidence that recognition of other emotions differed materially between modalities. These effects were found after taking into account age, estimated pre-morbid IQ, general facial recognition ability, an index of

executive skills and gender and so cannot be attributed to between-group differences in these abilities.

This work adds to a large body of evidence showing that facial emotion recognition is impaired in early HD. In common with a number of other studies, facial emotion recognition was impaired in the three negative emotions included in this paradigm. Many other studies have reported an impairment in facial fear and anger recognition (see e.g., Calder et al., 2010; Snowden et al., 2008 for recent examples). An impairment in facial sadness recognition has been found in about two thirds of the studies in which it was tested although this seems more variable and may to some extent be modified by the stimulus set and response options available (Snowden et al., 2008).

Far fewer studies have investigated vocal emotion recognition in HD, and in those that have, fear and anger recognition are often found to be impaired, but there have been no previous reports of an impairment in vocal sadness recognition (Calder et al., 2010; Hayes et al., 2007; Snowden et al., 2008; Sprengelmeyer et al., 1996). Even in the current study the impairment in vocal sadness recognition disappeared with the exclusion of the five participants who had done badly on the Montreal Battery of Evaluation of Amusia, which suggests that this effect was relatively weak, and may have been at least in part attributable to difficulties perceiving basic auditory qualities.

Emotion recognition in music, a more abstract modality, has not previously been examined in HD. Results mirrored those for the other two modalities, with impairments found in recognising angry and fearful music, and a suggestion that sad music recognition might be impaired although the latter effect was not found when

participants failing the Amusia Battery were excluded. Overall, therefore, this early HD cohort demonstrates a cross-modal deficit in the recognition of angry and fearful stimuli.

There has been much debate in the literature as to whether recognition of particular emotions is disproportionately impaired in HD. Whilst some early studies argued in favour of a disproportionate impairment in disgust recognition (both facial and vocal), later and larger studies have tended to demonstrate either that anger is worst affected, or that the magnitude of the impairment does not differ significantly between negative emotions (compare e.g., Calder et al., 2010; Gray et al., 1997; Henley et al., 2008; Johnson et al., 2007; Sprengelmeyer et al., 1996). In the current study participants tended to find fearful stimuli harder to recognise than angry or sad stimuli in the same modality. However, as disgusting stimuli were not included in this study, and the response options were limited to four instead of the more usual six, this finding simply suggests that any disproportionate impairment is likely to depend on the experimental paradigm and adds weight to the view that recognition of most negative emotions is impaired, to some extent, in HD.

Control performance in this study was relatively similar to that seen in other studies (e.g., Henley et al., 2008; Omar et al., 2011), with the group tending to perform well on most stimuli, but finding angry and fearful musical stimuli harder to recognise. It was noted that, on average, the control group mis-identified two of the forty emotion stimuli, reporting an angry face as "sad" and a fearful musical clip as "angry". This is not an unusual finding bearing in mind that control participants were deliberately chosen to share social backgrounds with HD participants; two previous studies have found impairments in facial anger recognition in gene-negative controls relative to

"normal" healthy controls (Gray et al., 1997; Sprengelmeyer, Schroeder, Young, & Epplen, 2006). As this is the first report of music emotion recognition in HD only future work will confirm whether mis-identification of fear in music is a chance finding or common in this group.

Whilst some previous work has included facial and vocal stimuli in the same battery, inter-modality performance has not been compared statistically in any previous study. In the current study recognition of fearful stimuli appeared to be easier for the HD group in the musical modality than in facial and vocal modalities. However there was no evidence of inter-modality differences for sad or angry stimuli, suggesting that the level of impairment for these two emotions is broadly similar regardless of modality. This has implications for the cognitive locus of impairment in HD; it seems likely that the impairment seen is at least partly attributable to an inability to access relatively high-level emotional conceptual or semantic knowledge, not just modality-specific instances of each emotion.

This conclusion is consistent with previous research suggesting that recognition of emotion in music uses a similar network of brain regions as recognition of emotion in other stimuli (e.g. ventral striatum, amygdala, prefrontal cortex, insula (see Koelsch, 2010 for a review)). Where components of this network are compromised we would predict cross-modal deficits, independent of the modality tested; modality-specific deficits might be more dependent on difficulties with particular perceptual processes unique to each modality. Indeed, other neurological populations have been shown to have impaired recognition of emotions across similar modalities, including those with Parkinson's disease (van Tricht et al., 2010) and fronto-temporal lobar degeneration (FTLD) (Omar et al., 2011; Snowden et al., 2008). In the former FTLD

population poor music emotion recognition performance was associated with reduced grey matter volume in a number of regions, including insular, anterior cingulate, orbitofrontal regions, hippocampus and amygdala, and the ventral striatum, all of which have been implicated in emotion processing in imaging studies of healthy participants, in facial, vocal and musical modalities (Adolphs, 2002). There was also evidence of reduced grey matter in other regions, such as the medial pre-frontal cortex, which might be unique to music-specific processes and are also implicated in skills such as theory of mind. It may be that the abstract nature of the musical stimulus requires participants to attribute a "mental state" to it, similar to what is required in theory of mind judgements, and that impairments in underlying brain structure impact on both sorts of task (Omar et al., 2011). However it is unclear whether a similar skills deficit might underlie the music emotion recognition impairment seen in the current HD cohort, as recent evidence suggests that their social cognition difficulties may be qualitatively different to those seen in FTLD, and not be attributable to a fundamental theory of mind deficit (Snowden et al., 2003).

4.2 LIMITATIONS

A major limitation of the current study was the fact that the musical stimuli consisted of excerpts from existing Western classical music. This made it impossible to control for familiarity or prior semantic association. Although participants were asked to indicate whether they thought a clip was familiar, responses suggested that this was not a reliable index; for example participants sometimes claimed to recognise a clip from a particular piece, but said that they had never heard a later clip taken from the same piece (which seems unlikely). Many clips will have been unfamiliar to many participants, forcing them to make a novel judgement about the emotion. However some clips will have been familiar, and in particular may have

been associated with emotive events such as weddings, which could well have influenced response choice. This is most likely to have introduced noise into these data, and it is not clear to what extent this was the case, or whether there might have been a systematic bias for some more popular stimuli. However, there is no evidence to suggest that familiarity would have differed systematically between control and HD participants, and hence it seems unlikely that it would have inflated group differences. A potential solution to these concerns is to compose novel stimuli. This has in fact been done, since the first set of data were collected for the current study, and a set of novel music emotion stimuli now exists (Hailstone et al., 2009).

A related issue is that musical stimuli tend to induce higher physiological arousal than faces or voices (with neural correlates including amygdala and related systems) and therefore this is a potential confound when comparing behavioural (and imaging) data between musical and non-musical modalities (Omar et al., 2011). Reduced arousal in neurological populations relative to controls may therefore inflate group differences in music emotion recognition scores. In future studies it should be possible to use some measure of autonomic arousal in order to control or at least adjust for this.

Participants reported a range of musical backgrounds (see Appendix 5), from those who rarely listened to music, to those who went about their daily lives with the radio on, and some who had had several years of formal musical training and still participated in amateur music-making. Again, whilst there was no apparent bias towards either controls or HD participants being relatively more "musical", it was not possible to quantify this in a meaningful way, and hence prior musical experience was not controlled for. However, there is evidence that lack of musical training does

not prevent listeners from making accurate emotional categorisations from music (Peretz et al., 1998), and that music emotion recognition is consistent across listeners with a variety of musical backgrounds (Krumhansl, 1997) so it seems unlikely that varying musical backgrounds had a large influence on the results reported here. Nevertheless an effort to match groups for some measure of musical background would be needed in future work in order to minimise doubts about this.

Finally, fewer than half the participants completed the Montreal Battery of Evaluation of Amusia (Peretz et al., 2003), meaning that it was not possible to adjust for this statistically in the main analysis. Of those who completed it similar (small) proportions of control and early HD participants failed, but those HD participants who failed did very poorly, possibly suggesting a qualitative difference between them and the controls. Reanalysing the data without these participants however did not change the major findings of impaired anger and fear recognition across modalities, but did suggest that the finding of impaired recognition of sad voices and music might have been attributable, at least in part, to aspects of musical perceptual ability, rather than emotion recognition skills *per se*. In future it would be important to test all participants on some measure of musical perception in order to assess the influence of these skills on auditory emotion recognition generally. There is increasing evidence that around 4% of the general population are "tone deaf", i.e. born without the ability to discriminate tunes in music (congenital amusia) (Peretz, Cummings, & Dube, 2007; Stewart, 2008) and the Peretz battery is one way of testing for this.

4.3 FUTURE WORK

Future work should initially focus on replicating the results reported here, as this is the first study to test music emotion recognition in HD, and to test statistically the differences between a number of modalities. As mentioned above, it is important to control for factors such as age, education, executive function and gender as was done here, but also to attempt to minimise possible between-group differences due to musical ability and background. Thus future studies would benefit from taking steps to do this, perhaps matching groups on some measure of "musicality", and / or ensuring stringent testing for basic musical perceptual abilities. In addition, the use of novel musical stimuli is recommended. Regarding the facial and vocal stimuli, there is also some debate as to the ecological validity of the Ekman stimuli, which although very widely-used and therefore easily comparable, are static and black-and-white, and therefore very unlike stimuli that would be encountered in everyday life. Some attempts to make moving facial stimuli have been made (Montagne et al., 2006; Simon, Craig, Gosselin, Belin, & Rainville, 2007) which as well as seeming more ecologically-valid, would be a better match for auditory stimuli which are necessarily temporally dynamic. Also, given that stimulus set can influence results within a single cohort (Snowden et al., 2008), it would seem important to conduct more studies in which more than one stimulus set is used in a given modality, in order to establish effects that are independent of this factor.

Given the debate around the neural substrate underlying emotion perception in general, and music perception in particular, more structural and functional imaging studies of multi-modal emotion recognition should be encouraged. In HD, as in FTLD, the pattern and progression of brain atrophy is relatively well characterised. Previous work in HD has used both voxel-based morphometry and functional

magnetic resonance imaging (fMRI) to examine neural correlates of facial emotion recognition in HD (Henley et al., 2008; Hennenlotter et al., 2004; Kipps et al., 2007) and this work could usefully be expanded to investigate which regions are implicated in emotion recognition across modalities in this population. Unfortunately it was not possible to conduct an imaging study on the current cohort as not all participants had structural MRIs during their research visits, and those that did were performed on different scanners, with different field strengths, introducing several potential confounds.

Comparatively few studies have looked at emotion recognition in presymptomatic HD (participants who carry the gene but are not yet judged to be showing unequivocal motor signs of the disease). In part this may be because it is often hard to quantify how far from predicted disease onset these cohorts are, and cognitive deficits can be very subtle prior to motor onset, and therefore hard to detect using standard clinical tests. Nevertheless with sensitive tests and adequate cohorts some changes can be detected many years prior to motor "onset" (Stout et al., 2011; Tabrizi et al., 2009) and it would be of interest to quantify the emotion recognition deficit better in this population, and to track how it changes over time. To date very few studies in presymptomatic or symptomatic HD have looked in detail at emotion recognition longitudinally (an exception being Sprengelmeyer et al., 2006).

Finally, there is very little published work examining the socio-emotional effects of this impairment on people with HD and their families and carers. One recent study reports that impaired facial emotion recognition is associated with impaired everyday function (Ille et al., 2011), and these authors, along with many others, hypothesise that a breakdown in affect recognition could well underlie other social

communication problems seen in HD. Others note that whilst people with HD recognise that their communication is slower and more effortful, their families and carers report that personality changes (such as difficulty changing perspective) affect communication negatively as the disease progresses (Hartelius, Jonsson, Rickeberg, & Laakso, 2010), and that interpersonal difficulties can be an early sign of the disease (Williams et al., 2007). However, there does not seem to be an empirical study designed to assess whether deficits in emotion recognition have a direct impact on social communication, or to examine carers' perspectives and experiences. An ability to recognise and ascribe mental states (theory of mind) is a key skill needed for empathy, and many of the brain areas implicated in theory of mind and pro-social behaviour overlap with those thought to be affected in HD (Singer, 2006). Therefore it would seem important that future studies address not only the nature of the deficits in HD, but the effect of this on those around people with HD. Increased understanding of the causes of the social communication problems in the disease may help both patients and carers manage these difficulties more successfully, and thus go some way to mitigating their effects, even when the symptoms themselves cannot be alleviated.

4.4 CONCLUSION

The emotion recognition impairment in HD appears to be cross-modal, with recognition of angry and fearful stimuli in particular equally affected in the facial, vocal and musical domains. This study does not support suggestions that particular emotions or modalities are disproportionately affected in early HD. This suggests the involvement of relatively high-level neural systems, rather than modality-specific mechanisms. In terms of the social impact of the disease, this means that unfortunately efforts to substitute information from one modality with that from

another are unlikely to mitigate the effects of impairment. Further work should focus on using more ecologically-valid stimuli, and following up presymptomatic cohorts over time in order further to clarify the trajectory of the emotion recognition deficit in this population.

5 REFERENCES

Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr.Opin.Neurobiol., 12,* 169-177.

Aviezer, H., Bentin, S., Hassin, R. R., Meschino, W. S., Kennedy, J., Grewal, S. et al. (2009). Not on the face alone: perception of contextualized face expressions in Huntington's disease. *Brain, 132,* 1633-1644.

Benton, A. L., Hamsher, K. S., Varney, N., & Spreen, O. (1983). *Contributions to neuropsychological assessment: A clinical manual.* Oxford: Oxford University Press.

Blood, A. J. & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc.Natl.Acad.Sci.U.S.A, 98,* 11818-11823.

Brown, S., Martinez, M. J., & Parsons, L. M. (2004). Passive music listening spontaneously engages limbic and paralimbic systems. *Neuroreport, 15,* 2033-2037.

Calder, A. J., Keane, J., Lawrence, A. D., & Manes, F. (2004). Impaired recognition of anger following damage to the ventral striatum. *Brain, 127,* 1958-1969.

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nat.Neurosci, 3,* 1077-1078.

Calder, A. J., Keane, J., Young, A. W., Lawrence, A. D., Mason, S., & Barker, R. A. (2010). The relation between anger and different forms of disgust:

implications for emotion recognition impairments in Huntington's disease. *Neuropsychologia, 48,* 2719-2729.

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nat.Rev.Neurosci., 2,* 352-363.

Cohen, J. (1992). A power primer. *Psychol.Bull., 112,* 155-159.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage, 55,* 705-712.

Ekman, P. & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P. (1992). Are there basic emotions? *Psychol Rev., 99,* 550-553.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav.Res.Methods, 39,* 175-191.

Gray, J. M., Young, A. W., Barker, W. A., Curtis, A., & Gibson, D. (1997). Impaired recognition of disgust in Huntington's disease gene carriers. *Brain, 120,* 2029-2038.

Griffiths, T. D., Warren, J. D., Dean, J. L., & Howard, D. (2004). "When the feeling's gone": a selective loss of musical emotion. *J.Neurol.Neurosurg.Psychiatry, 75,* 344-345.

Hailstone, J. C., Omar, R., Henley, S. M., Frost, C., Kenward, M. G., & Warren, J. D. (2009). It's not what you play, it's how you play it: timbre affects

perception of emotion in music. *Q.J.Exp.Psychol.(Colchester.), 62,* 2141-2155.

Harper, P. (2002). The epidemiology of Huntington's disease. In G.Bates, P. Harper, & L. Jones (Eds.), *Huntington's Disease* (3 ed., pp. 3-27). Oxford: Oxford University Press.

Hartelius, L., Jonsson, M., Rickeberg, A., & Laakso, K. (2010). Communication and Huntington's disease: qualitative interviews and focus groups with persons with Huntington's disease, family members, and carers. *Int.J.Lang Commun.Disord., 45,* 381-393.

Hayes, C. J., Stevenson, R. J., & Coltheart, M. (2007). Disgust and Huntington's disease. *Neuropsychologia, 45,* 1135-1151.

Henley, S. M., Wild, E. J., Hobbs, N. Z., Frost, C., MacManus, D. G., Barker, R. A. et al. (2009). Whole-brain atrophy as a measure of progression in premanifest and early Huntington's disease. *Mov Disord., 24,* 932-936.

Henley, S. M., Wild, E. J., Hobbs, N. Z., Warren, J. D., Frost, C., Scahill, R. I. et al. (2008). Defective emotion recognition in early HD is neuropsychologically and anatomically generic. *Neuropsychologia, 46,* 2152-2160.

Henley, S. M. D., Novak, M. J. U., Frost, C., King, J., Tabrizi, S. J., & Warren, J. D. (2011). Emotion recognition in Huntington's disease: a systematic review. *Neuroscience and Biobehavioral Reviews*, in press

Hennenlotter, A., Schroeder, U., Erhard, P., Haslinger, B., Stahl, R., Weindl, A. et al. (2004). Neural correlates associated with impaired disgust processing in pre-symptomatic Huntington's disease. *Brain, 127,* 1446-1453.

Hsieh, F. Y., Bloch, D. A., & Larsen, M. D. (1998). A simple method of sample size calculation for linear and logistic regression. *Stat.Med., 17,* 1623-1634.

Huntington Study Group (1996). Unified Huntington's Disease Rating Scale: reliability and consistency. Huntington Study Group. *Mov Disord, 11,* 136-142.

Ille, R., Holl, A. K., Kapfhammer, H. P., Reisinger, K., Schafer, A., & Schienle, A. (2011). Emotion recognition and experience in Huntington's disease: Is there a differential impairment? *Psychiatry Res.,* In press.

Johnson, S. A., Stout, J. C., Solomon, A. C., Langbehn, D. R., Aylward, E. H., Cruce, C. B. et al. (2007). Beyond disgust: impaired recognition of negative emotions prior to diagnosis in Huntington's disease. *Brain, 130,* 1732-1744.

Kipps, C. M., Duggins, A. J., McCusker, E. A., & Calder, A. J. (2007). Disgust and happiness recognition correlate with anteroventral insula and amygdala volume respectively in preclinical Huntington's disease. *J Cogn Neurosci, 19,* 1206-1217.

Koelsch, S. (2010). Towards a neural basis of music-evoked emotions. *Trends Cogn Sci., 14,* 131-137.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Can.J.Exp.Psychol., 51,* 336-353.

Lawrence, A. D., Hodges, J. R., Rosser, A. E., Kershaw, A., ffrench, C. C., Rubinsztein, D. C. et al. (1998). Evidence for specific cognitive deficits in preclinical Huntington's disease. *Brain, 121,* 1329-1341.

Milders, M., Crawford, J. R., Lamb, A., & Simpson, S. A. (2003). Differential deficits in expression recognition in gene-carriers and patients with Huntington's disease. *Neuropsychologia, 41,* 1484-1492.

Mitchell, I. J., Heims, H., Neville, E. A., & Rickards, H. (2005). Huntington's disease patients show impaired perception of disgust in the gustatory and olfactory modalities. *J Neuropsychiatry Clin Neurosci, 17,* 119-121.

Montagne, B., Kessels, R. P., Kammers, M. P., Kingma, E., de Haan, E. H., Roos, R. A. et al. (2006). Perception of Emotional Facial Expressions at Different Intensities in Early-Symptomatic Huntington's Disease. *Eur Neurol, 55,* 151-154.

Nelson, H. E. & Willison, J. (1991). *The National Adult Reading Test*. (2 ed.) Windsor, UK: NFER-Nelson.

Omar, R., Hailstone, J. C., Warren, J. E., Crutch, S. J., & Warren, J. D. (2010). The cognitive organization of music knowledge: a clinical analysis. *Brain, 133,* 1200-1213.

Omar, R., Henley, S. M., Bartlett, J. W., Hailstone, J. C., Gordon, E., Sauter, D. A. et al. (2011). The structural neuroanatomy of music emotion recognition: Evidence from frontotemporal lobar degeneration. *Neuroimage, 56,* 1814-1821.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Ann.N.Y.Acad.Sci., 999,* 58-75.

Peretz, I., Cummings, S., & Dube, M. P. (2007). The genetics of congenital amusia (tone deafness): a family-aggregation study. *Am.J.Hum.Genet., 81,* 582-588.

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition, 68,* 111-141.

Reitan, R. M. & Wolfson, D. (2004). *Comprehensive Handbook of Psychological Assessment. Volume 1. Intellectual and Neuropsychological Assessment.* Hoboken, NJ: John Wiley and Sons.

Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology, 1,* 43-46.

Sauter, D. (2006). *An investigation into vocal expressions of emotions: the roles of valence, culture, and acoustic factors.* Unpublished PhD Thesis, University College London.

Sauter, D. A. & Eimer, M. (2010). Rapid detection of emotion from human vocalizations. *J.Cogn Neurosci., 22,* 474-481.

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc.Natl.Acad.Sci.U.S.A, 107,* 2408-2412.

Shoulson, I. & Fahn, S. (1979). Huntington disease: clinical care and evaluation. *Neurology, 29,* 1-3.

Simon, D., Craig, K. D., Gosselin, F., Belin, P., & Rainville, P. (2007). Recognition and discrimination of prototypical dynamic expressions of pain and emotions. *Pain, 135,* 55-64.

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neurosci.Biobehav.Rev., 30,* 855-863.

Snowden, J. S., Austin, N. A., Sembi, S., Thompson, J. C., Craufurd, D., & Neary, D. (2008). Emotion recognition in Huntington's disease and frontotemporal dementia. *Neuropsychologia, 46,* 2638-2649.

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

Snowden, J. S., Gibbons, Z. C., Blackshaw, A., Doubleday, E., Thompson, J., Craufurd, D. et al. (2003). Social cognition in frontotemporal dementia and Huntington's disease. *Neuropsychologia, 41,* 688-701.

Speedie, L. J., Brake, N., Folstein, S. E., Bowers, D., & Heilman, K. M. (1990). Comprehension of prosody in Huntington's disease. *J Neurol Neurosurg Psychiatry, 53,* 607-610.

Sprengelmeyer, R., Schroeder, U., Young, A. W., & Epplen, J. T. (2006). Disgust in pre-clinical Huntington's disease: A longitudinal study. *Neuropsychologia, 44,* 518-533.

Sprengelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., Homberg, V. et al. (1996). Loss of disgust - perception of faces and emotions in Huntington's disease. *Brain, 119,* 1647-1665.

Stewart, L. (2008). Fractionating the musical mind: insights from congenital amusia. *Curr.Opin.Neurobiol., 18,* 127-130.

Stout, J. C., Paulsen, J. S., Queller, S., Solomon, A. C., Whitlock, K. B., Campbell, J. C. et al. (2011). Neurocognitive signs in prodromal Huntington disease. *Neuropsychology, 25,* 1-14.

Tabrizi, S. J., Langbehn, D. R., Leavitt, B. R., Roos, R. A., Durr, A., Craufurd, D. et al. (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.

van Tricht, M. J., Smeding, H. M., Speelman, J. D., & Schmand, B. A. (2010). Impaired emotion recognition in music in Parkinson's disease. *Brain Cogn, 74,* 58-65.

Williams, J. K., Hamilton, R., Nehl, C., Gonigal-Kenney, M., Schutte, D. L., Sparbel, K. et al. (2007). "No one else sees the difference: "family members' perceptions of changes in persons with preclinical Huntington disease. *Am.J.Med.Genet.B Neuropsychiatr.Genet., 144,* 636-641.

PART 3: CRITICAL APPRAISAL

1 INTRODUCTION

This appraisal reflects on three main aspects of conducting the research reported in Parts 1 and 2. Firstly, there is a brief discussion of working on the systematic review, the pitfalls encountered and how these would be better managed in future. Secondly, it expands on the methodological limitations of the empirical work discussed briefly in Part 2. Finally, I reflect more on the personal impact of working in a research setting with participants with a neurodegenerative disease.

2 LITERATURE REVIEW

The review of the literature covered a set of studies reporting quantitative results of emotion recognition performance in HD, a subject about which there is much debate. When planning the review, it was thought that a meta-analytic approach would be a sensible way to analyse such data. A meta-analysis would potentially answer questions that have arisen about whether some emotions are disproportionately affected, and also allow consideration of the relationship between demographic variables and emotion recognition performance across cohorts. The main difficulty with this approach was the unavailability of much of the necessary raw data. Quantitative outcomes were not reported in the published papers for several studies, and in addition many studies did not report the necessary demographic data such as IQ, CAG repeat length and UHDRS motor score. Attempts were made to contact representatives of all the included studies. However not all these attempts were successful, and even those that were sometimes resulted in data being promised, but not being sent before the review deadline, or in authors admitting that they had never had CAG repeat data and could not access their results a few years after the original study was performed. On consideration it seemed that attempting a meta-analysis on

an incomplete data set would potentially raise as many questions as it answered, and subsequently the review was changed to be entirely narrative.

If undertaking such a review again I would approach this problem differently. Firstly, with the experience of having done one such review, I would be quicker to identify what data I needed and what data were missing from the included studies. I would then contact authors much earlier in the review process, and hence have more time to chase unanswered queries as the process moved on. In general the responses I received were as helpful as they could be, and also many people expressed an interest in the work I was doing. This was encouraging and also useful to have made links with other people in the field.

3 EMPIRICAL STUDY

3.1 RECRUITMENT

Having had previous experience of research I was aware that recruitment was likely to take a long time, and that I would need to allocate a number of research days to making telephone calls, writing letters, and the administrative side of the project. With this in mind I submitted my ethics application as early as possible with a view to beginning to contact potential participants in the middle of 2010 and finishing data collection by the end of 2010. This worked well and I was able to collect data on 20 out of 25 of my extra participants within the planned timescale. Unfortunately one of my last controls was unable to be visited because of unexpected bad weather, and then work commitments meant he had to withdraw from the study. This left me trying to recruit one more participant in early January 2011, and ultimately meant data collection continued sporadically until early February. On reflection, I should have anticipated more problems with recruiting from this population, as although

they tend to be very keen to be involved in research, there are a lot of projects running concurrently at the National Hospital which means that some participants are too busy to take on more. However, overall because I had planned ahead with this aspect, data collection did not overrun by too much time.

3.2 METHODOLOGICAL CHOICES

In putting together the test battery there were a number of considerations. It needed to be brief enough to administer in a single visit, well-tolerated, and the materials needed to be easily transportable as I was planning on testing people in their homes. The multimodal emotion test had already been developed and was the key test of interest, but background measures were needed for potential confounding variables such as estimated pre-morbid IQ, face recognition and musical skills, and executive function which is known to be affected in HD, and was thought might impact on emotion recognition performance. Standard measures of IQ, face recognition and musical perception were included.

3.2.1 Executive function

There is a wide range of executive function tests and the Trail-Making Test was chosen partly because it is brief to administer and relatively free from floor and ceiling effects in the early HD population, and also because subtracting time A from time B yields an index of "task switching" with the motor component removed. With hindsight, I might consider using an alternative test such as phonemic fluency, as this is even simpler to administer and still an index of "frontal" lobe function.

3.2.2 Musical perception

The Peretz battery was very long, and verged on being intolerable for some participants. Participants must listen to 120 pairs of repeated tunes, and identify whether there is a one-note difference between each pair. Most reported finding this hard, and losing concentration. I tried to overcome this by encouraging them to stay focused, reminding them that they could request to hear stimuli again, and also breaking the test up (doing the first two sets of 30, and then having a short pause before completing the second two sets). However the effects of reduced concentration and motivation cannot be ruled out on this test, and were I to use it in future studies I would investigate whether a briefer version could be developed and normed, or whether one or two subtests would provide sufficient data rather than four. An addition problem in the current study is that not everyone completed the Peretz battery (as discussed in Part 2), meaning that it could not be included as a confounding variable in the main analysis. Were I to conduct further follow-up studies I would certainly ensure that all participants had done some measure of music perception.

3.2.3 Musical emotion recognition

The multimodal battery was developed using existing facial and non-verbal vocal stimuli, with the musical extracts added by me. This was conceived as a pilot study, for which it was simpler to use existing classical music than to construct novel stimuli. The extracts used were tested extensively prior to inclusion in the final battery, and the modal control response in the current study matched that of the target emotion for all stimuli except one. This suggests that despite variability, stimuli were generally consistently identified with the target emotion. The one exception was a fearful stimulus which 10/25 controls identified as this, whilst 13/25 classed it as

angry, and 2/25 as sad. However, as reported in Parts 1 and 2, controls who are gene negative but related to someone carrying the HD gene are known to have slightly impaired emotion recognition skills relative to other healthy volunteers, so it is possible that this difference represented a characteristic of this control group, rather than a problem with the stimulus *per se*.

One clear drawback of this approach is that people may have prior semantic links with certain clips of Western classical music, for example if they have heard them on television, or at a wedding. Informally, participants were asked to indicate whether or not a clip was familiar, but this proved to be subjective and unreliable, with for example participants hearing two clips from the same piece of music and indicating that one was familiar whilst the other was not; or spontaneously mis-identifying a clip (e.g. labelling Mozart's overture to The Marriage of Figaro as Beethoven's Fifth Symphony, or stating "that's not Jaws" on hearing the theme from Jaws). Thus it was not possible to include familiarity judgements as a covariate, to see whether familiarity affected emotional attribution, as these judgements were clearly extremely unreliable and likely to simply add noise to the analysis.

A more rigorous approach would obviously have been to compose novel stimuli, something which could not be done in the time available for this project. However a set of novel musical stimuli have now been made (Hailstone et al., 2009) and the promising results of the current study suggest that further research with these novel stimuli may be of value.

Another difficultly with this approach was that in order to be consistent between modalities, disgusting and surprising stimuli were omitted as it was considered impossible to render them musically. In some ways this says something about the

nature of those two emotions, particularly disgust which, when portrayed facially or vocally tends to imply a visceral distaste for something physical; the more subtle meaning of moral disapproval does not seem to come across as well in typical examples of facial or vocal expressions (Calder et al., 2010). Nevertheless, most other studies of emotion recognition in HD include these two emotions as stimuli and of course response options. Different emotions are differentially confusable (disgust with anger, and fear with surprise, in particular) (Calder et al., 1996), which means that excluding some materially changes the nature of the test (and can lead to different conclusions using the same group of participants (Snowden et al., 2008)) and that is a potential weakness of this study.

3.2.4 Participants

The multidisciplinary HD clinic at the National Hospital in London is one of the largest clinics in the country, and there are very good links between clinical and research teams, meaning that there is a large body of potential participants. Many people with HD prefer not to get involved in research, but there are also many who are keen to be involved in as much as possible. There are currently several large, international longitudinal studies running at the National Hospital (e.g. Track-HD, Tabrizi et al., 2009), which, despite the willingness of patients to become involved in research, means that there is a risk that patients become over-tested, or that studies might begin to confound each other. For example, it is unwise for cognitive testing to be repeated within a six-month period (and preferably longer) because of practice effects (see e.g., Bachoud-Lévi et al., 2001). This means that when starting a relatively small, single-visit study such as this, there is a limit to potential recruits. Currently many of the working-age patients with HD are enrolled in Track-HD, which involves a whole day of testing annually, but may lead on to potential

treatment trials in the future. Because of the timescale of testing for my study, I could only enrol participants from Track-HD if their Track-HD test dates were at least six months away from my test date. In addition, because of the time commitment involved in Track-HD many participants could not commit themselves to further research studies. What this tended to mean was that I had a disproportionate number of participants at either end of the age spectrum; some who were too young for Track-HD, and some who were past retirement age and therefore freer to attend a number of research visits. With more time, and perhaps support from other researchers, I would have been able to balance this, but within the timescale available to me this was not possible.

3.2.5 Including imaging data

Given the debate in the literature about emotion recognition in HD and other domains, it would have been a useful addition to have included some structural imaging in the study. The original 25 participants had structural MRI scans as part of the London Longitudinal study. When originally planning the extension it was thought that if I could recruit the extra 25 participants from existing imaging studies I could use their structural MRIs (with permission). Whilst every participant consented to this, unfortunately it was not possible to recruit all 25 from existing imaging studies, and that meant that several participants did not have imaging available. In addition there would have been two other potentially large confounds: the original 25 were scanned on a 1.5T scanner, whilst later participants were scanned on a 3T scanner; and the original 25 had scans on the same day as cognitive testing, whereas the later participants who had imaging had often had it several months before I saw them. Time and importantly financial constraints meant that it

was not possible to include a structural MRI scan in the protocol for this study, and therefore the imaging part could not be completed.

3.3 WORKING WITH PEOPLE WITH A NEURODEGENERATIVE DISEASE

I have worked with people with HD since 2003, and therefore came to this study with a good idea of what was involved, particularly emotionally, in meeting people whose lives are being devastated by the disease. Even so, there were times when it was particularly hard to manage my feelings, at the same time as maintaining a professional outlook and balancing the needs of the study with the needs of the participants.

Some of the participants whom I tested in the last 12 months were people whom I had first met seven years ago. In some cases their disease seemed very little advanced, and it was encouraging to see them and hear how they had managed. In other cases people had very obviously deteriorated, as could be seen by their motor symptoms and also in their accounts, and their partners' accounts, of difficulties with everyday tasks (such as managing finances) and personal relationships. It was very sad to experience this, and it was hard, and sometimes impossible, to "switch off" at the end of a testing session and not reflect on the changes I had noticed.

There were also situations with participants whom I hadn't met before which were challenging. One family had only recently had HD diagnosed in an elderly participant, and were still adjusting to the realisation that either or both of their children might be affected, and their children's children, and how to manage this. Mine was the first research study they had done, and understandably they were full of hope that research might provide a cure for the next generation, but also full of questions. I had to try to be honest in my answers, acknowledging the limits of my

knowledge even when there was pressure just to reassure, and also acknowledging the limits of my role – as researcher, not clinician.

Another participant was one of the youngest, yet had relatively advanced disease. This participant had two young children, and the non-gene-carrying partner spoke briefly of the struggles to cope with family life, and with the affected partner's nearby family, who were finding it hard to come to terms with the diagnosis. I had visited them at home, so could see for myself the difficulties they were both facing, and I could imagine the challenges that would present as the disease got worse. Again, it was very difficult to focus on research in the face of this, and to ensure that I spent some time hearing what the participants wanted to say to me, and reflecting their concerns, as well as helping them complete the cognitive tests.

It was also not just participants with HD who presented challenges such as this. One young control was married to an at-risk person, and this was the first research study she had enrolled in. She broke down at the beginning of testing, talking about the difficulties of living with uncertainty, and was so distressed that I spent some time completing a risk assessment with her, and making sure she was aware of external support systems that she could access if needed. Ultimately she opted to complete the testing, because she felt that this was the only way she could gain some control over the disease, and "give something back", which is a common feeling amongst most of the people I tested.

It was while conducting research with this population before that I became frustrated with my inability to offer them any clinical help, and decided to apply for clinical training. In a way it was frustrating to find myself back in that position, being careful not to blur the line between researcher and clinician. However that was

balanced by the fact that I do have many more clinical skills now than a few years ago, which meant that even within the research setting I felt more confident about taking some time just to listen, reflect back dilemmas to participants, assess risk, and discuss whether they would like to be put in touch with other agencies that could offer more clinical help. I also had both my supervisors to report to, as well as the support network of the HD clinical and research teams at the National Hospital, and lay agencies such as the UK Huntington's Disease Association (www.hda.org.uk).

Overall, the emotional impact of working with this population was challenging, and sometimes led me to question the "usefulness" of what I was doing, particularly when contrasted with my day-to-day clinical work in which I could more obviously apply my skills to helping people with mental distress. However, all the participants knew, as I did, that this study was not about changing their quality of life, but simply about shedding more light on the nature of the disease. They were still willing to participate, and nearly all expressed their belief that by doing so they were helping, and fighting back. That attitude in itself is one of the reasons why it is also extremely rewarding, and inspirational to work with this population.

4 SUMMARY

Overall, whilst this study has methodological limitations as outlined above, I think it was successful in terms of piloting a new concept (music emotion recognition) with the HD population. Reflecting on the limitations has been useful in clarifying what changes should be made in future studies, and given that this is the first time music emotion stimuli have been used in HD, it would be wise to try to replicate and expand on the findings presented here.

Working with people with HD and their families also presents personal challenges from both a clinical and research point of view. However, their willingness to get involved, and their hope for the future, mean that it is also a rewarding and motivating experience.

5 REFERENCES

Bachoud-Lévi, A. C., Maison, P., Bartolomeo, P., Boisse, M. F., Dalla, B. G., Ergis, A. M. et al. (2001). Retest effects and cognitive decline in longitudinal follow-up of patients with early HD. *Neurology, 56,* 1052-1058.

Calder, A. J., Keane, J., Young, A. W., Lawrence, A. D., Mason, S., & Barker, R. A. (2010). The relation between anger and different forms of disgust: implications for emotion recognition impairments in Huntington's disease. *Neuropsychologia, 48,* 2719-2729.

Calder, A. J., Young, A. W., Rowland, D., Perrett, D. I., Hodges, J. R., & Etcoff, N. L. (1996). Facial emotion recognition after bilateral amygdala damage: Differentially severe impairment of fear. *Cognitive Neuropsychology, 13,* 699-745.

Hailstone, J. C., Omar, R., Henley, S. M., Frost, C., Kenward, M. G., & Warren, J. D. (2009). It's not what you play, it's how you play it: timbre affects perception of emotion in music. *Q.J.Exp.Psychol.(Colchester.), 62,* 2141-2155.

Snowden, J. S., Austin, N. A., Sembi, S., Thompson, J. C., Craufurd, D., & Neary, D. (2008). Emotion recognition in Huntington's disease and frontotemporal dementia. *Neuropsychologia, 46,* 2638-2649.

Tabrizi, S. J., Langbehn, D. R., Leavitt, B. R., Roos, R. A., Durr, A., Craufurd, D. et al. (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.

APPENDIX 1: ARTICLES THAT DID NOT MEET SEARCH CRITERIA FOR LITERATURE REVIEW

Reference	Reason for exclusion
Nabilone shows potential for symptomatic relief of Huntington's disease (2008). *Pharmacy in Practice, 18,* 183.	RCT for drug treatment, No emotion recognition tasks in battery.
Abel, C. G., Stein, G., Arakaki, T., Mancuso, M., Nano, G., Garretto, N. *et al.* (2007). Decision making ability assessment in patients with basal ganglia and cerebellum subcortical syndromes: Parkinson, Huntington and isolated degenerative cerebellar diseases. [Spanish]. *Revista Neurologica Argentina, 32,* 20-34.	Includes social cognition tests but they are Theory of Mind and gambling, not emotion recognition.
Arango-Lasprilla, J. C., Rogers, H., Lengenfelder, J., Deluca, J., Moreno, S., & Lopera, F. (2006). Cortical and subcortical diseases: do true neuropsychological differences exist? *Arch.Clin.Neuropsychol., 21,* 29-40.	No emotion recognition tasks in battery.
Bachoud-Levi, A. C., Maison, P., Bartolomeo, P., Boisse, M. F., Dalla, B. G., Ergis, A. M. *et al.* (2001). Retest effects and cognitive decline in longitudinal follow-up of patients with early HD. *Neurology, 56,* 1052-1058.	No emotion recognition tasks in battery.
Baker, J. G. (1996). Memory and emotion processing in cortical and subcortical dementia. *J.Gen.Psychol., 123,* 185-191.	Summary of old research, nothing novel.
Bales, K. R. (2004). Neurodegenerative disease research in the 21st century. *Drug Discovery Today, 9,* 553-556.	Conference review.
Bamford, K. A., Caine, E. D., Kido, D. K., Cox, C., & Shoulson, I. (1995). A prospective evaluation of cognitive decline in early Huntington's disease: functional and radiographic correlates. *Neurology, 45,* 1867-1873.	No emotion recognition tasks in battery.
Barquero-Jimenez, M. S. & Gomez-Tortosa, E. (2001). [Cognitive disorders in patients with Huntington's disease]. *Rev.Neurol., 32,* 1067-1071.	No emotion recognition tasks in battery.
Baudic, S., Maison, P., Dolbeau, G., Boisse, M. F., Bartolomeo, P., Dalla, B. G. *et al.* (2006). Cognitive impairment related to apathy in early Huntington's disease. *Dement.Geriatr.Cogn Disord., 21,* 316-321.	No emotion recognition tasks in battery.
Berrios, G. E., Wagle, A. C., Markova, I. S., Wagle, S. A., Rosser, A., & Hodges, J. R. (2002). Psychiatric symptoms in neurologically asymptomatic Huntington's disease gene carriers: a comparison with gene negative at risk subjects. *Acta Psychiatr.Scand., 105,* 224-230.	No emotion recognition tasks in battery.
Blackmore, L., Simpson, S. A., & Crawford, J. R. (1995). Cognitive performance in UK sample of presymptomatic people carrying the gene for Huntington's disease. *J.Med.Genet., 32,* 358-362.	No emotion recognition tasks in battery.
Bodner, T., Jenner, C., Benke, T., Ober, A., Seppi, K., & Fleischhacker, W. W. (2001). Intoxication with riluzole in Huntington's disease. *Neurology, 57,* 1141-1143.	Case report.
Bonelli, R. M. & Kapfhammer, H. P. (2003). Why minocycline is helpful in Huntington's disease. *J.Psychopharmacol., 17,* 461.	Drug report.
Boxer, A. L. & Yoon, G. (2007). Reply from the authors [6]. *Neurology, 68,* 1325.	Reply to query about juvenile HD.

Reference	Reason for exclusion
Brandt, J., Inscore, A. B., Ward, J., Shpritz, B., Rosenblatt, A., Margolis, R. L. *et al.* (2008). Neuropsychological deficits in Huntington's disease gene carriers and correlates of early "conversion". *J.Neuropsychiatry Clin.Neurosci., 20,* 466-472.	No emotion recognition tasks in battery.
Brandt, J., Leroi, I., O'Hearn, E., Rosenblatt, A., & Margolis, R. L. (2004). Cognitive impairments in cerebellar degeneration: a comparison with Huntington's disease. *J.Neuropsychiatry Clin.Neurosci., 16,* 176-184.	No emotion recognition tasks in battery.
Brandt, J., Shpritz, B., Codori, A. M., Margolis, R., & Rosenblatt, A. (2002). Neuropsychological manifestations of the genetic mutation for Huntington's disease in presymptomatic individuals. *J.Int.Neuropsychol.Soc., 8,* 918-924.	No emotion recognition tasks in battery.
Campodonico, J. R., Codori, A. M., & Brandt, J. (1996). Neuropsychological stability over two years in asymptomatic carriers of the Huntington's disease mutation. *J.Neurol.Neurosurg.Psychiatry, 61,* 621-624.	No emotion recognition tasks in battery.
de Boo, G. M., Tibben, A., Lanser, J. B., Jennekens-Schinkel, A., Hermans, J., Maat-Kievit, A. *et al.* (1997). Early cognitive and motor symptoms in identified carriers of the gene for Huntington disease. *Arch.Neurol., 54,* 1353-1357.	No emotion recognition tasks in battery.
de Gelder, B., Van den Stock, J., Balaguer, R. D., & Bachoud-Levi, A. C. (2008). Huntington's disease impairs recognition of angry and instrumental body language. *Neuropsychologia, 46,* 369-373.	Emotion matching (within modality), not explicit emotion recogntion, labelling or cross-modality matching.
Derouesne, C. (2004). [Cognitive disorders at the onset of Huntington disease]. *Psychol.Neuropsychiatr.Vieil., 2,* 226-227.	Editorial, no novel data.
Duff, K., Beglinger, L. J., Theriault, D., Allison, J., & Paulsen, J. S. (2010). Cognitive deficits in Huntington's disease on the Repeatable Battery for the Assessment of Neuropsychological Status. *J.Clin.Exp.Neuropsychol., 32,* 231-238.	No emotion recognition tasks in battery.
Fletcher, L. (1997). Computer 'games' diagnose early Huntington's disease. *Mol.Med.Today, 3,* 48-49.	Focuses on CANTAB as diagnostic tool for HD, not emotion recognition.
Giordani, B., Berent, S., Boivin, M. J., Penney, J. B., Lehtinen, S., Markel, D. S. *et al.* (1995). Longitudinal neuropsychological and genetic linkage analysis of persons at risk for Huntington's disease. *Arch.Neurol., 52,* 59-64.	No emotion recognition tasks in battery.
Gomez-Anson, B., Alegret, M., Munoz, E., Monte, G. C., Alayrach, E., Sanchez, A. *et al.* (2009). Prefrontal cortex volume reduction on MRI in preclinical Huntington's disease relates to visuomotor performance and CAG number. *Parkinsonism Relat Disord., 15,* 213-219.	No emotion recognition tasks in battery.
Gomez-Anson, B., Alegret, M., Munoz, E., Sainz, A., Monte, G. C., & Tolosa, E. (2007). Decreased frontal choline and neuropsychological performance in preclinical Huntington disease. *Neurology, 68,* 906-910.	No emotion recognition tasks in battery.
Gomez-Tortosa, E., del, B. A., Garcia Ruiz, P. J., Pernaute, R. S., Benitez, J., Barroso, A. *et al.* (1998). Severity of cognitive impairment in juvenile and late-onset Huntington disease. *Arch.Neurol., 55,* 835-843.	No emotion recognition tasks in battery.
Hahn-Barma, V., Deweer, B., Durr, A., Dode, C., Feingold, J., Pillon, B. *et al.* (1998). Are cognitive changes the first symptoms of Huntington's disease? A study of gene carriers. *J.Neurol.Neurosurg.Psychiatry, 64,* 172-177.	No emotion recognition tasks in battery.
Halligan, P. W. (1998). Inability to recognise disgust in Huntington's disease. *Lancet, 351,* 464.	Commentary, no novel data.

Reference	Reason for exclusion
Hoth, K. F., Paulsen, J. S., Moser, D. J., Tranel, D., Clark, L. A., & Bechara, A. (2007). Patients with Huntington's disease have impaired awareness of cognitive, emotional, and functional abilities. *J.Clin.Exp.Neuropsychol., 29,* 365-376.	No emotion recognition tasks in battery.
Jacobs, D. H., Shuren, J., & Heilman, K. M. (1995). Impaired perception of facial identity and facial affect in Huntington's disease. *Neurology, 45,* 1217-1218.	Includes emotion testing but tests matching and discriminating, not labelling / explicit recognition; also no controls.
Jason, G. W., Suchowersky, O., Pajurkova, E. M., Graham, L., Klimek, M. L., Garber, A. T. *et al.* (1997). Cognitive manifestations of Huntington disease in relation to genetic structure and clinical onset. *Arch.Neurol., 54,* 1081-1088.	No emotion recognition tasks in battery.
Jurgens, C. K., van de, W. L., van Es, A. C., Grimbergen, Y. M., Witjes-Ane, M. N., Van Der, G. J. *et al.* (2008). Basal ganglia volume and clinical correlates in 'preclinical' Huntington's disease. *J.Neurol., 255,* 1785-1791.	No emotion recognition tasks in battery.
Lawrence, A. D., Hodges, J. R., Rosser, A. E., Kershaw, A., Ffrench-Constant, C., Rubinsztein, D. C. *et al.* (1998a). Evidence for specific cognitive deficits in preclinical Huntington's disease. *Brain, 121,* 1329-1341.	No emotion recognition tasks in battery.
Lawrence, A. D., Watkins, L. H., Sahakian, B. J., Hodges, J. R., & Robbins, T. W. (2000). Visual object and visuospatial cognition in Huntington's disease: implications for information processing in corticostriatal circuits. *Brain, 123,* 1349-1364.	No emotion recognition tasks in battery.
Lawrence, A. D., Weeks, R. A., Brooks, D. J., Andrews, T. C., Watkins, L. H., Harding, A. E. *et al.* (1998b). The relationship between striatal dopamine receptor binding and cognitive performance in Huntington's disease. *Brain, 121,* 1343-1355.	No emotion recognition tasks in battery.
Lemiere, J., Decruyenaere, M., Evers-Kiebooms, G., Vandenbussche, E., & Dom, R. (2002). Longitudinal study evaluating neuropsychological changes in so-called asymptomatic carriers of the Huntington's disease mutation after 1 year. *Acta Neurol.Scand., 106,* 131-141.	No emotion recognition tasks in battery.
Lemiere, J., Decruyenaere, M., Evers-Kiebooms, G., Vandenbussche, E., & Dom, R. (2004). Cognitive changes in patients with Huntington's disease (HD) and asymptomatic carriers of the HD mutation--a longitudinal follow-up study. *J.Neurol., 251,* 935-942.	No emotion recognition tasks in battery.
Lichter, D. G. & Hershey, L. A. (2010). Before chorea. Pre-Huntington mild cognitive impairment. *Neurology.75*, 490-491.	Commentary, no novel data.
Morris, M. (1995). Dementia and cognitive changes in Huntington's disease. *Adv.Neurol., 65,* 187-200.	Review, no data.
Nehl, C. & Paulsen, J. S. (2004). Cognitive and psychiatric aspects of Huntington disease contribute to functional capacity. *J.Nerv.Ment.Dis., 192,* 72-74.	No emotion recognition tasks in battery.
Paulsen, J. S. & Conybeare, R. A. (2005). Cognitive changes in Huntington's disease. *Adv.Neurol., 96,* 209-225.	Review, no data.

Reference	Reason for exclusion
Paulsen, J. S., Hayden, M., Stout, J. C., Langbehn, D. R., Aylward, E., Ross, C. A. *et al.* (2006). Preparing for preventive clinical trials: the Predict-HD study. *Arch.Neurol., 63,* 883-890.	Same data are presented (but with slightly more participants) in the Johnson *et al.* 2007 paper which is included in review (Julie Stout, personal communication)
Paulsen, J. S., Langbehn, D. R., Stout, J. C., Aylward, E., Ross, C. A., Nance, M. *et al.* (2008). Detection of Huntington's disease decades before diagnosis: the Predict-HD study. *J.Neurol.Neurosurg.Psychiatry, 79,* 874-880.	No emotion recognition tasks in battery.
Paulsen, J. S., Wang, C., Duff, K., Barker, R., Nance, M., Beglinger, L. *et al.* (2010). Challenges assessing clinical endpoints in early Huntington disease. *Mov Disord., 25,* 2595-2603	No emotion recognition tasks in battery.
Pierrot-Deseilligny, C. (2001). Actualites American Academy of Neurology Philadelphie, 5-11 mai 2001* compte-rendu du congres. [French]. *Revue Neurologique, 157,* 578-600.	Congress account, no data.
Pillon, B., Dubois, B., & Agid, Y. (1996). Testing cognition may contribute to the diagnosis of movement disorders. *Neurology, 46,* 329-334.	Review, no data.
Redondo, V. L., Brown, R. G., & Chacon, J. (2001). [Executive dysfunction in Huntington's disease]. *Rev.Neurol., 32,* 923-929.	No emotion recognition tasks in battery.
Robins Wahlin, T. B., Lundin, A., & Dear, K. (2007). Early cognitive deficits in Swedish gene carriers of Huntington's disease. *Neuropsychology, 21,* 31-44.	No emotion recognition tasks in battery.
Rodrigues, G. R., Souza, C. P., Cetlin, R. S., de Oliveira, D. S., Pena-Pereira, M., Ujikawa, L. T. *et al.* (2009). Use of the frontal assessment battery in evaluating executive dysfunction in patients with Huntington's disease. *J.Neurol., 256,* 1809-1815.	No emotion recognition tasks in battery.
Roger, K. S. (2005). Exploring memory loss: A study starts. *Journal of Dementia Care, 13,* 36.	Commentary.
Rogers, D. (1993). Movement disorders. *Current Opinion in Psychiatry, 6,* 113-116.	Review, no data.
Roitberg, B. (2004). Research news and notes. *Surgical Neurology, 61,* 106-108.	Commentary.
Rosas, H. D., Salat, D. H., Lee, S. Y., Zaleta, A. K., Pappu, V., Fischl, B. *et al.* (2008). Cerebral cortex and the clinical expression of Huntington's disease: complexity and heterogeneity. *Brain., 131,* 1057-1068.	No emotion recognition tasks in battery.
Rosas, H. D., Tuch, D. S., Hevelone, N. D., Zaleta, A. K., Vangel, M., Hersch, S. M. *et al.* (2006). Diffusion tensor imaging in presymptomatic and early Huntington's disease: Selective white matter pathology and its relationship to clinical measures. *Mov Disord., 21,* 1317-1325.	No emotion recognition tasks in battery.
Rosenberg, N. K., Sorensen, S. A., & Christensen, A. L. (1995). Neuropsychological characteristics of Huntington's disease carriers: a double blind study. *J.Med.Genet., 32,* 600-604.	No emotion recognition tasks in battery.

Reference	Reason for exclusion
Rupp, J., Blekher, T., Jackson, J., Beristain, X., Marshall, J., Hui, S. *et al.* (2010). Progression in prediagnostic Huntington disease. *J.Neurol.Neurosurg.Psychiatry, 81,* 379-384.	No emotion recognition tasks in battery.
Sawa, A. & Snyder, S. H. (2005). Two genes link two distinct psychoses. *Science, 310,* 1128-1129.	Genes for psychosis.
Sax, D. S., Powsner, R., Kim, A., Tilak, S., Bhatia, R., Cupples, L. A. *et al.* (1996). Evidence of cortical metabolic dysfunction in early Huntington's disease by single-photon-emission computed tomography. *Mov Disord., 11,* 671-677.	No emotion recognition tasks in battery.
Simpson, S. A. (2004). The management of Huntington's disease. *Practical Neurology, 4,* 204-211.	Review, no data.
Snowden, J. S., Craufurd, D., Thompson, J., & Neary, D. (2002). Psychomotor, executive, and memory function in preclinical Huntington's disease. *J.Clin.Exp.Neuropsychol., 24,* 133-145.	No emotion recognition tasks in battery.
Soliveri, P., Monza, D., Piacentini, S., Paridi, D., Nespolo, C., Gellera, C. *et al.* (2002). Cognitive and psychiatric characterization of patients with Huntington's disease and their at-risk relatives. *Neurol.Sci., 23 Suppl 2,* S105-S106.	No emotion recognition tasks in battery.
Solomon, A. C., Stout, J. C., Weaver, M., Queller, S., Tomusk, A., Whitlock, K. B. *et al.* (2008). Ten-year rate of longitudinal change in neurocognitive and motor function in prediagnosis Huntington disease. *Mov Disord., 23,* 1830-1836.	No emotion recognition tasks in battery.
Sprengelmeyer, R., Young, A. W., Sprenglemeyer, A., Calder, A. J., Rowland, D., Perrett, D. *et al.* (1997). Recognition of facial expressions: Selective impairment of specific emotions in Huntington's disease. *Cognitive Neuropsychology, Vol.14,* 839-879.	Two case studies.
Sprengelmeyer, R. (2007). The neurology of disgust. *Brain, 130,* 1715-1717.	Commentary.
Stout, J. C., Weaver, M., Solomon, A. C., Queller, S., Hui, S., Johnson, S. A. *et al.* (2007). Are cognitive changes progressive in prediagnostic HD? *Cogn Behav.Neurol., 20,* 212-218.	No emotion recognition tasks in battery.
Thieben, M. J., Duggins, A. J., Good, C. D., Gomes, L., Mahant, N., Richards, F. *et al.* (2002). The distribution of structural neuropathology in pre-clinical Huntington's disease. *Brain., 125,* 1815-1828.	No emotion recognition tasks in battery.
Thompson, J. C., Poliakoff, E., Sollom, A. C., Howard, E., Craufurd, D., & Snowden, J. S. (2010). Automaticity and attention in Huntington's disease: when two hands are not better than one. *Neuropsychologia, 48,* 171-178.	No emotion recognition tasks in battery.
Thompson, J. C., Snowden, J. S., Craufurd, D., & Neary, D. (2002). Behavior in Huntington's disease: dissociating cognition-based and mood-based changes. *J.Neuropsychiatry Clin.Neurosci., 14,* 37-43.	No emotion recognition tasks in battery.
Timman, R., Tibben, A., & Roos, R. A. (2003). Nonlinear effects in behavioral changes in Huntington disease. *Cogn Behav.Neurol., 16,* 82.	Letter, no novel data.
Tost, H., Wendt, C. S., Schmitt, A., Heinz, A., & Braus, D. F. (2004). Huntington's disease: phenomenological diversity of a neuropsychiatric condition that challenges traditional concepts in neurology and psychiatry. *Am.J.Psychiatry, 161,* 28-34.	Case study.
van Oostrom, J. C., Dekker, M., Willemsen, A. T., de Jong, B. M., Roos, R. A., & Leenders, K. L. (2009). Changes in striatal dopamine D2 receptor binding in pre-clinical Huntington's disease. *Eur.J.Neurol., 16,* 226-231.	No emotion recognition tasks in battery.

Reference	Reason for exclusion
van Walsem, M. R., Sundet, K., Retterstol, L., & Sundseth, O. (2010). A double blind evaluation of cognitive decline in a Norwegian cohort of asymptomatic carriers of Huntington's disease. *J.Clin.Exp.Neuropsychol., 32,* 590-598.	No emotion recognition tasks in battery.
Verny, C., Allain, P., Prudean, A., Malinge, M. C., Gohier, B., Scherer, C. *et al.* (2007). Cognitive changes in asymptomatic carriers of the Huntington disease mutation gene. *Eur.J.Neurol., 14,* 1344-1350.	No emotion recognition tasks in battery.
Veyssier-Belot, C. (2005). Psychiatric disorders and systemic diseases. [French]. *Revue de Medecine Interne, 26,* 682-685.	Review, no novel data.
Videnovic, A., Bernard, B., Fan, W., Jaglin, J., Leurgans, S., & Shannon, K. M. (2010). The Montreal Cognitive Assessment as a screening tool for cognitive dysfunction in Huntington's disease. *Mov Disord., 25,* 401-404.	No emotion recognition tasks in battery.
Wang, K., Hoosain, R., Yang, R. M., Meng, Y., & Wang, C. Q. (2003). Impairment of recognition of disgust in Chinese with Huntington's or Wilson's disease. *Neuropsychologia, 41,* 527-537.	Genetic confirmation not available (Wang, personal communication).
Ward, J., Sheppard, J. M., Shpritz, B., Margolis, R. L., Rosenblatt, A., & Brandt, J. (2006). A four-year prospective study of cognitive functioning in Huntington's disease. *J.Int.Neuropsychol.Soc., 12,* 445-454.	No emotion recognition tasks in battery.
Wetter, S., Peavy, G., Jacobson, M., Hamilton, J., Salmon, D., & Murphy, C. (2005). Olfactory and auditory event-related potentials in Huntington's disease. *Neuropsychology, 19,* 428-436.	Investigates odour perception but not emotion or disgust.
Wetter, S. R. (2003). *Olfactory psychophysics and electrophysiology in huntington's disease.*	Thesis - based on same concepts as Wetter *et al.* (2005), above.
Wexler, A. (2006). Huntington disease [2]. *Journal of the Royal Society of Medicine, 99,* 53.	Letter, no novel data.
Wild, E. J. & Tabrizi, S. J. (2006). Predict-HD and the future of therapeutic trials. *Lancet Neurology, 5,* 724-725.	Commentary.
Wilkinson, D. & Halligan, P. (2004). The relevance of behavioural measures for functional-imaging studies of cognition. *Nat.Rev.Neurosci., 5,* 67-73.	Commentary.
Williams, R. (2006). Hunting for huntingtin modification. *Nature Reviews Neuroscience, 7,* 503.	Research highlights, no novel data.
Witjes-Ane, M. N., Mertens, B., van Vugt, J. P., Bachoud-Levi, A. C., van Ommen, G. J., & Roos, R. A. (2007). Longitudinal evaluation of "presymptomatic" carriers of Huntington's disease. *J.Neuropsychiatry Clin.Neurosci., 19,* 310-317.	No emotion recognition tasks in battery.
Witjes-Ane, M. N., Vegter-Van, D., V, van Vugt, J. P., Lanser, J. B., Hermans, J., Zwinderman, A. H. *et al.* (2003). Cognitive and motor functioning in gene carriers for Huntington's disease: a baseline study. *J.Neuropsychiatry Clin.Neurosci., 15,* 7-16.	No emotion recognition tasks in battery.
Young, A. W., Sprengelmeyer, R., Phillips, M., & Calder, A. J. (1997). Response from Young, Sprengelmeyer, Phillips and Calder. [References]. *Trends in Cognitive Sciences, Vol.1,* 322-325.	Response to comments about original Sprengelmeyer paper, no novel data.
Zakzanis, K. K. (1998). The subcortical dementia of Huntington's disease. *J.Clin.Exp.Neuropsychol., 20,* 565-578.	Review, no novel data.
Zihl, J. (2004). Clear indications of emotion depend on vivid stimuli. *Journal of Neurology, Neurosurgery and Psychiatry,75,* 1658-1659.	Comment on emotion recognition testing.

APPENDIX 2: INDIVIDUAL STUDIES (FROM INCLUDED REFERENCES) THAT DID NOT MEET SEARCH CRITERIA FOR LITERATURE REVIEW

Study and reference	Reason for exclusion
Aviezer *et al.* (2009): Experiment 2	Participants were asked to identify facial emotion expressions that were superimposed on body images portraying a different emotion and thus the experiment was not assessing "pure" emotion recognition in either modality.
Calder *et al.* (2010): Study 2	Participants were asked to match photographs of different "types" of disgust recognition with a written scenario, i.e. was too specific for this review.
Hayes *et al.* (2007): Experiments 1, 3, 4, 5, 6, 7	These experiments did not include an overt recognition component. Expt. 1 required participants to describe situations that would induce emotions, in Expt. 3 participants had to categorise emotional words, in Expt. 4 they had to categorise emotion-inducing scenes, all of which might rely solely or in part on semantic knowledge. In Expt. 5 their experience of disgust was assessed using a questionnaire, and in Expts. 6 and 7 they were asked to rate odours and tastes but this did not include the explicit label "disgust".
Milders *et al.* (2003): Test 2	Test 2 involved matching facial expressions and therefore did not meet criteria for explicit emotion recognition.
Snowden *et al.* (2008): Tasks 1, 2, 3, 6, 7	Task 1 required participants to define emotion labels, Task 2 asked participants to pick synonyms or link emotion words with specific scenarios, and Task 3 repeated Task 2 but with reduced response options; thus these tasks were not examining explicition recognition of emotional stimuli. Task 6 was a facial expression matching task, and Task 7 assessed facial identity matching.

APPENDIX 3: ETHICAL APPROVAL AND RELATED PAPERWORK

1. Ethical approval letter
2. Letter to potential participants
3. Control consent form
4. Patient consent form
5. Control information sheet
6. Patient information sheet

ETHICAL APPROVAL LETTER

The National Hospital for Neurology and Neurosurgery & Institute of Neurology Joint REC
REC Office
South House
Royal Free Hospital
Pond Street
London
NW3 2QG
Tel: 020 7794 0500 ext. 31342
Fax: 0207 7941004
Email:
Sasha.Vandayar@royalfree.nhs.uk
Website: www.uclh.nhs.uk

[redacted]

Our Ref 010L 135

25 March 2010

Dear Dr Henley

Study Title:	**Multimodal emotion recognition in Huntington's Disease**
REC reference number:	**10/H0716/9**
Protocol number:	**2**

Thank you for your letter of 08 March 2010, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information was considered by a sub-committee of the REC at a meeting held on 24/03/2010 A list of the sub-committee members is attached.

Confirmation of ethical opinion
On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Ethical review of research sites
The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion
The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

For NHS research sites only, management permission for research ("R&D approval") should be obtained from the relevant care organisation(s) in accordance

with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk. *Where the only involvement of the NHS organisation is as a Participant Identification Centre, management permission for research is not required but the R&D office should be notified of the study. Guidance should be sought from the R&D office where necessary.*

Sponsors are not required to notify the Committee of approvals from host organisations.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Covering Letter	1	06 January 2010
REC application	2.5	08 January 2010
Protocol	2	14 February 2009
Letter of invitation to participant	1	14 December 2009
Referees or other scientific critique report	1	23 October 2009
Questionnaire: Hailstone Music Questionnaire	1	15 September 2009
GP/Consultant Information Sheets	1	14 December 2009
Response to Request for Further Information		08 March 2010
Participant Information Sheet: Controls	4	25 February 2010
Participant Information Sheet: Patients	4	25 February 2010
Participant Consent Form: Controls	3	25 February 2010
Participant Consent Form: Patients	3	25 February 2010

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

10/H0716/9	**Please quote this number on all correspondence**

Yours sincerely

Dr Yogi Amin

Chair

Email: sasha.vandayar@royalfree.nhs.uk

Enclosures: List of names and professions of members who were present at the meeting and those who submitted written comments [if final opinion was confirmed was given at a meeting]

"After ethical review – guidance for researchers" [SL-AR1 for CTIMPs, SL- AR2 for other studies]

Copy to: Philip Diamond
R&D office for NHS care organisation at lead site

The National Hospital for Neurology and Neurosurgery & Institute of Neurology Joint REC

Attendance at Sub-Committee of the REC meeting on 24 March 2010

Committee Members:

Name	*Profession*	*Present*	*Notes*
Dr Yogi Amin	Consultant in Neuroanaesthesia & Neurocritical Care	Yes	
Dr Lorraine Ludman	Chartered Psychologist	Yes	

The National Hospital for Neurology and Neurosurgery
Queen Square
London WC1N 3BG
Telephone: 020 7837 3611

LETTER TO POTENTIAL PARTICIPANTS

PRIVATE AND CONFIDENTIAL
Name
Address

Date

Dear

Following our recent telephone conversation I am writing to send you further information about the research study "Emotion recognition in Huntington's disease". I am enclosing an information sheet which contains details about why we are doing the study, and what is involved. Please take your time to read it and think about whether or not you would be interested in taking part. You might find it useful to discuss it with other people or to telephone me if you have any questions.

As I said on the phone, I will wait for a few weeks for you to ring me and let me know whether you would like to take part in the study. If I do not hear from you for more than a month I will ring you to check whether or not you are still interested.

If you decide to take part, then when we next speak on the telephone I will arrange a time to meet up with you. When we meet I will check whether you have any more questions and ask you to sign a consent form before you do the tasks. We can either do them at your home, or you can come into London and do them at UCL (in Queen Square near the National Hospital), whichever you prefer.

If you decide not to take part that is fine, and once I am aware of that I will not contact you any more about the study.

Yours sincerely,

Dr Susie Henley
Trainee Clinical Psychologist
e-mail: susie.henley@ucl.ac.uk
phone: [redacted]

CONTROL CONSENT FORM

Name of study: Emotion recognition in Huntington's disease
Name of researcher: Dr Susie Henley

Please initial each box

Study information I confirm that I have read and understood the information sheet dated 25/02/2010 v 4.0 for the above study. I have had time to consider the information, ask questions, and have had these answered satisfactorily.	☐
Medical records I give my permission for members of the research team to view my medical records.	☐
MRI scan If I have previously had an MRI scan for research purposes at UCL/UCLH, I give permission for this scan to be used in the current study.	☐
Data protection and Data Sharing I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.	☐
Voluntary participation I understand that my participation is entirely voluntary. I am free to withdraw at any time, without giving a reason, and without my legal rights or medical care being affected.	☐
Contacting GP I consent to your contacting my GP to inform him / her of my participation in this study.	☐
Agreement I agree to take part in this study.	☐

Name of participant	Date	Signature
Name of person taking consent	Date	Signature

When completed, 1 for participant; 1 for researcher site file; 1 (original) to be kept in medical notes.

PATIENT CONSENT FORM

Name of study: Emotion recognition in Huntington's disease
Name of researcher: Dr Susie Henley

Please initial each box

Study information I confirm that I have read and understood the information sheet dated 25/02/2010 v 4.0 for the above study. I have had time to consider the information, ask questions, and have had these answered satisfactorily.	☐
Medical records I give my permission for members of the research team to view my medical records.	☐
MRI scan If I have previously had an MRI scan for research purposes at UCL/UCLH, I give permission for this scan to be used in the current study.	☐
Data protection and Data Sharing I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.	☐
Voluntary participation I understand that my participation is entirely voluntary. I am free to withdraw at any time, without giving a reason, and without my legal rights or medical care being affected.	☐
Contacting GP I consent to your contacting my GP to inform him / her of my participation in this study.	☐
Agreement I agree to take part in this study.	☐

Name of participant	Date	Signature
Name of person taking consent	Date	Signature

When completed, 1 for participant; 1 for researcher site file; 1 (original) to be kept in medical notes.

The National Hospital for Neurology and Neurosurgery
Queen Square
London WC1N 3BG
Telephone: 020 7837 3611

CONTROL INFORMATION SHEET

Name of study: Emotion recognition in Huntington's disease
We would like to invite you to take part in a research study. Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Talk to others about the study if you wish.

Part 1 tells you the purpose of this study and what will happen to you if you take part.

Part 2 gives you more detailed information about the conduct of the study.

Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

PART 1
What is the purpose of the study?
We already know that Huntington's disease (HD) can affect how people recognise emotions. For example, people with HD sometimes have difficulty recognising faces that look angry, or disgusted, or frightened. The main question of this study is whether the difficulties people with HD have with emotion recognition mainly affects faces, or whether they show similar difficulties recognising other kinds of emotion, in voices and music (in other words, a more general problem understanding emotion rather than just a problem with facial expressions). We hope that this will help us understand a little more about what it is like to have HD, and which areas of the brain are affected.

Why have I been invited to take part?
You have been invited because you are a partner, spouse or carer of someone with HD, or you were at risk of inheriting HD but have had a negative test for the gene mutation. We are hoping that a total of 25 people with HD will take part in the study, as well as 25 people without HD.

Do I have to take part?
You do not have to take part in the study. It is up to you to decide, and it is voluntary, which means that you can change your mind at any point, even after the study has started. If you choose not to take part that is fine and will not affect any future medical care in any way. If you would like to take part we will ask you to sign a consent form, but you can still decide to leave the study whenever you want.

In order to take part in the study you must not meet any of the following criteria:

- Aged under 18 or over 65
- Past head injury
- Past or current excessive use of alcohol
- Past or current drug use
- Current medical, neurological or mental health issue
- English is not your first language

If you meet any of these criteria you do not need to disclose details to the researcher, but please let the researcher know that you do not wish to take part in the study.

What will happen to me if I take part?
If you would like to take part you will be asked to spend about two or three hours doing some tasks with a research psychologist. This includes a questionnaire about your musical abilities but it does not matter whether you are musical or not (it is not a test of musical ability). It also includes some pencil and paper tasks. You will also be asked to look and listen to some pictures and sounds presented on a computer, but you do not need to be able to use a computer in order to answer the questions.

There will be time to take breaks and have refreshments if you need them.

You only need to do the tasks once, and this can be in your own home, or at the National Hospital in London, depending on what you would prefer. We can refund second-class travel expenses for you if you opt to travel to the hospital to do the tasks.

If you have had a brain scan (a magnetic resonance imaging or MRI scan) as part of another study, for example Track-HD, or the London HD study, then we will ask your permission to use that scan for this study. **You do not have to have a brain scan for this study**.

Are there any risks or side-effects?
There are no physical risks from doing the tasks. As this is a research study you will not be told how you have done on the tasks.

What are the possible benefits of taking part?
As the study does not involve any treatments, taking part will not help you medically. We hope that the results will tell us more about what it is like to have HD which may help us give more advice to people with HD and their carers.

What if there is a problem?
Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

Will my taking part be kept confidential?
Yes. We will follow the principles of good research governance in line with the Data Protection Act, and all information about you will be handled in confidence. The details are included in Part 2. We will ask your permission to inform your GP of your participation in the study.

PART 2
What will happen if I decide I want to leave the study half-way through?
If you decide you want to leave the study at any point that is fine. We will delete any research records we have for you and any results we have for you will not be used in the study.

How will my data be stored?
All of the data we collect about you is confidential and stored within the terms of the Data Protection Act (1998).

The results are given a code unique to you (we do not use your real name). The results are kept securely at UCL and only the people running and supervising the

study have access to them. People from regulatory authorities can also request to have access to the data in order to check that all the results are being stored properly. These people will also keep your data confidential.

Your data are only being used for this study.

We need to keep the data for about five years after we publish the results, because sometimes after results have been published people have questions about the findings. After this time the records are destroyed.

What will happen to the results of the study?
After the results are analysed we will write to you with a summary of the findings. We will also try to publish the results in academic journals, and we may present them at conferences. As we are interested in group results, we will never present or discuss your individual results with anyone. You will never be identified in any report.

Who is organising and funding the study?
The study is organised by Dr Susie Henley, working with two supervisors, Dr John King (University College London) and Dr Jason Warren (Institute of Neurology). Susie is undertaking the study as part of her doctoral course in Clinical Psychology.

Who has reviewed the study?
All research in the NHS is looked at by independent group of people, called a Research Ethics Committee, to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by the National Hospital for Neurology and Neurosurgery Research Ethics Committee.

Complaints
If you have a concern about any aspect of this study, you should ask to speak first to Dr Henley who will do her best to answer your questions (07528 254982). If you remain unhappy and wish to complain formally, you can do this through the NHS Complaints Procedure. Details can be obtained from the hospital.

Compensation arrangements
If you are harmed by taking part in this research project, there are no special compensation arrangements. If you are harmed due to someone's negligence, then you may have grounds for a legal action but you may have to pay for it. Regardless of this, if you wish to complain, or have any concerns of this study, the normal National Health Service complaints mechanisms should be available to you.

Who can I contact for more information?
You may contact Dr Susie Henley on [redacted] or susie.henley@ucl.ac.uk

You will be given a copy of this information sheet to keep.

The National Hospital for Neurology and Neurosurgery
Queen Square
London WC1N 3BG
Telephone: 020 7837 3611

PATIENT INFORMATION SHEET

Name of study: Emotion recognition in Huntington's disease

We would like to invite you to take part in a research study. Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Talk to others about the study if you wish.

Part 1 tells you the purpose of this study and what will happen to you if you take part.

Part 2 gives you more detailed information about the conduct of the study.

Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

PART 1

What is the purpose of the study?

We already know that Huntington's disease (HD) can affect how people recognise emotions. For example, people with HD sometimes have difficulty recognising faces that look angry, or disgusted, or frightened. The main question of this study is whether the difficulties people with HD have with emotion recognition mainly affects faces, or whether they show similar difficulties recognising other kinds of emotion, in voices and music (in other words, a more general problem understanding emotion rather than just a problem with facial expressions). We hope that this will help us understand a little more about what it is like to have HD, and which areas of the brain are affected.

Why have I been invited to take part?

You have been invited because you have HD, and you have told us that you are interested in taking part in research. We are hoping that a total of 25 people with HD will take part in the study, as well as 25 people without HD.

Do I have to take part?

You do not have to take part in the study. It is up to you to decide, and it is voluntary, which means that you can change your mind at any point, even after the study has started. If you choose not to take part that is fine and will not affect any future medical care in any way. If you would like to take part we will ask you to sign a consent form, but you can still decide to leave the study whenever you want.

Taking part in a research study is quite separate from the care you receive in clinic and from your GP. Whether or not you take part in this study your medical care will carry on as normal and not be affected in any way.

In order to take part in the study you must not meet any of the following criteria:

- Aged under 18 or over 65
- Past head injury
- Past or current excessive use of alcohol
- Past or current drug use

- Current medical, neurological or mental health issue
- English is not your first language

If you meet any of these criteria you do not need to disclose details to the researcher, but please let the researcher know that you do not wish to take part in the study.

What will happen to me if I take part?
If you would like to take part you will be asked to spend about two or three hours doing some tasks with a research psychologist. This includes a questionnaire about your musical abilities but it does not matter whether you are musical or not (it is not a test of musical ability). It also includes some pencil and paper tasks. You will also be asked to look and listen to some pictures and sounds presented on a computer, but you do not need to be able to use a computer in order to answer the questions.

There will be time to take breaks and have refreshments if you need them.

You only need to do the tasks once, and this can be in your own home, or at the National Hospital in London, depending on what you would prefer. We can refund second-class travel expenses for you if you opt to travel to the hospital to do the tasks.

If you have had a brain scan (a magnetic resonance imaging or MRI scan) as part of another study, for example Track-HD, or the London HD study, then we will ask your permission to use that scan for this study. **You do not have to have a brain scan for this study**.

Are there any risks or side-effects?
There are no physical risks from doing the tasks. As this is a research study you will not be told how you have done on the tasks.

What are the possible benefits of taking part?
As the study does not involve any treatments, taking part will not help you medically. We hope that the results will tell us more about what it is like to have HD which may help us give more advice to people with HD and their carers.

What if there is a problem?
Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

Will my taking part be kept confidential?
Yes. We will follow the principles of good research governance in line with the Data Protection Act, and all information about you will be handled in confidence. The details are included in Part 2. We will ask your permission to inform your GP of your participation in the study.

PART 2
What will happen if I decide I want to leave the study half-way through?
If you decide you want to leave the study at any point that is fine. We will delete any research records we have for you and any results we have for you will not be used in the study.

How will my data be stored?
All of the data we collect about you is confidential and stored within the terms of the Data Protection Act (1998).

The results are given a code unique to you (we do not use your real name). The results are kept securely at UCL and only the people running and supervising the study have access to them. People from regulatory authorities can also request to have access to the data in order to check that all the results are being stored properly. These people will also keep your data confidential.

Your data are only being used for this study.

We need to keep the data for about five years after we publish the results, because sometimes after results have been published people have questions about the findings. After this time the records are destroyed.

What will happen to the results of the study?
After the results are analysed we will write to you with a summary of the findings. We will also try to publish the results in academic journals, and we may present them at conferences. As we are interested in group results, we will never present or discuss your individual results with anyone. You will never be identified in any report.

Who is organising and funding the study?
The study is organised by Dr Susie Henley, working with two supervisors, Dr John King (University College London) and Dr Jason Warren (Institute of Neurology). Susie is undertaking the study as part of her doctoral course in Clinical Psychology.

Who has reviewed the study?
All research in the NHS is looked at by independent group of people, called a Research Ethics Committee, to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by the National Hospital for Neurology and Neurosurgery Research Ethics Committee.

Complaints
If you have a concern about any aspect of this study, you should ask to speak first to Dr Henley who will do her best to answer your questions (07528 254982). If you remain unhappy and wish to complain formally, you can do this through the NHS Complaints Procedure. Details can be obtained from the hospital.

Compensation arrangements
If you are harmed by taking part in this research project, there are no special compensation arrangements. If you are harmed due to someone's negligence, then you may have grounds for a legal action but you may have to pay for it. Regardless of this, if you wish to complain, or have any concerns of this study, the normal National Health Service complaints mechanisms should be available to you.

Who can I contact for more information?
You may contact Dr Susie Henley on 07528 254982 or susie.henley@ucl.ac.uk

You will be given a copy of this information sheet to keep.

APPENDIX 4: HAILSTONE MUSIC QUESTIONNAIRE

1. Have you ever played a musical instrument? YES ☐ NO ☐

Instrument	Length of time in training (years)	Length of time playing (years)	Standard reached (grade level or equiv.)

2. Do you play an instrument or sing regularly? YES ☐ NO ☐

Instrument	Time playing per week (hours)	Where played (home/group)

3. Do you listen to music regularly? YES ☐ NO ☐

If Yes: (a) How many hours per week do you listen to music? _____

(b) What kind of music do you mainly listen to (pop, easy listening, jazz, classical, etc)

_____

_____

4. Tunefulness

Do you sing in tune? YES ☐ NO ☐

Does it sound in tune to you? YES ☐ NO ☐ CAN'T TELL ☐

Do other people think you sound in tune? YES ☐ NO ☐

5. Any other details that might be relevant?

_____
_____
_____
_____

APPENDIX 5: MUSICAL BACKGROUND OF PARTICIPANTS

The Hailstone questionnaire collected data on the number of years for which participants had played an instrument (including voice), and how many hours a week they currently listened to any genre of music. As both these variables were skewed within both groups, medians rather than means are presented. Median time playing was 1 year (range 0 – 20 years) in controls (n=24), and 2 years (range 0 – 30 years) in the HD group (n=21). Median time listening was 7 hours per week (range 0 – 70) in controls, and 7 hours per week (range 0 – 65) in the HD group. (Note as mentioned in Part 2, section 2.3.5, 1 control and 4 HD participants were not asked to complete this questionnaire as it was not available when they were assessed).

As the data were skewed and violated assumptions of Normality, Mann-Whitney U tests were used to compare playing time and listening time between groups. There was no evidence that years of playing differed statistically significantly between groups (z=-1.08, p=0.28) or that hours per week listening differed statistically significantly between groups (z=0.24, p=0.81).

Thus overall there was no evidence that in these measures of musical background, the groups differed significantly. However these variables could not be used as covariates in the main analysis, firstly because of missing data, and secondly because they are retrospective, variable and not necessarily equivalent between participants (e.g. 10 years' playing may equate to much more in one participant than another). These limitations are discussed further in Parts 2 and 3.

APPENDIX 6: PILOT STUDY OF MUSIC EMOTION STIMULI

Selection of stimuli was based on an initial pilot study in 16 healthy participants who did not participate in any subsequent experiments. Pilot participants were presented with a larger set of 104 musical excerpts (chosen by the author from Western classical music examples) and asked to rate each excerpt for how strongly it represented each of the four target emotions using a Likert scale ranging from 0 (not at all) to 4 (very strongly). Ratings for each excerpt for each emotion were averaged. An excerpt for which one and only one emotion achieved a mean rating ≥ 2 was considered to portray that emotion (other excerpts were considered insufficiently salient, or ambiguous). Excerpts fulfilling this criterion were ordered based on rating, and the 10 highest-ranking excerpts for each emotion were used in the test battery. Mean (range) ratings for each emotion were as follows: anger, 3.0 (2.8 – 3.8); fear, 3.1 (2.5 – 3.8); happiness, 3.2 (2.6 – 3.9); sadness, 2.8 (2.1 – 3.5).

Anger
Beethoven: Egmont Overture
Beethoven: Symphony No. 6: Storm
Dvorak: New World Symphony: Allegro con fuoco
Dvorak: New World Symphony: Scherzo
Elgar: Enigma Variation No. 4
Holst: The Planets: Mars
Saint Saens: Organ Symphony: Scherzo
Shostakovich: Symphony No. 5: Allegro
Shostakovich: Symphony No. 5: Moderato
Vivaldi: The Four Seasons: Summer

Fear
Bartok: Bluebeard's Castle: Lake of Tears
Bartok: Music for Strings, Percussion and Celesta
Goldenthal: Alien 3 Theme
Hermmann: Psycho Theme
Holst: The Planets: Saturn
Horner: Aliens Theme
Mussorgsky: Night on a Bare Mountain
Mussorgsky: Pictures at an Exhibition: Cum Mortuis
Schnittke: Concerto Grosso No. 3: Pesante
Williams: Jaws Theme

Happiness
Brahms: Romanze in F
Moross: Big Country Theme
Mozart: Marriage of Figaro Overture
Pachelbel: Canon in D
Puccini: La Boheme Overture
Rimsky-Korsakov: Capriccio Espagnol Alborada
Rimsky-Korsakov: Capriccio Espagnol Fandango
Smetana: Má Vlast - Vltava
Vivaldi: The Four Seasons: Autumn
Williams: Jurassic Park Theme

Sadness
Barber: Adagio for Strings
Beethoven: Pathetique Sonata: Grave
Brahms: Intermezzo in A major, Opus 118
Brahms: Symphony No 3: Poco Allegretto
Puccini: La Boheme Finale
Rimsky-Korsakov: Easter Festival Overture
Rimsky-Korsakov: Scheherezade
Shostakovich: Symphony No. 5: Allegro non troppo, clip 3
Vaughan Williams: Fantasia on a Theme by Thomas Tallis
Williams: Schindler's List Theme

APPENDIX 7: UNADJUSTED BEHAVIOURAL RESULTS FROM PART 2

Mean (SD) neuropsychological performance, with differences (95% confidence intervals) with and without adjustment for age and estimated premorbid IQ (see Part 2, section 3.2)

	Control	**HD**	**Difference (HD – Control)**	
	(N=25)	(N=25)	Crude	Adjusted
Benton Facial Recognition Test (/54)	48.4 (3.0)	45.5 (4.8)	-3.0 (-5.3, -0.7) p=0.013	
Trail-Making Test A, sec	23.3 (6.4)	38.6 (15.5)	15.3 (10.3, 23,2) p<0.05	14.2 (9.3, 21.4) p<0.05
Trail-Making Test B, sec	64.6 (29.0)	144.4 (80.4)	79.8 (49.1, 114.3) p<0.05	70.9 (43.9, 108.8) p<0.05
Trail-Making Test B – A, sec	41.2 (25.6)	105.7 (71.5)	64.5 (36.0, 94.1) p<0.05	56.7 (29.5, 87.4) p<0.05
Peretz Total Score (/120)	104.1 (8.7)	89.5 (11.0)	-14.6 (-23.4, -5.7) p=0.003	-12.8 (-22.2, -3.4) p=0.011

Note that the Benton score is already adjusted for age and education
Adjusted differences are assumed to be constant at all levels of the covariates adjusted for

Odds of a correct response (95% confidence intervals) relative to controls for each emotion in each modality, unadjusted (see Part 2, Table 3-2 for comparison)

	Intended emotion			
Faces	Happy	Sad	Angry	Fearful
Controls		1	1	1
Early HD		0.16 (0.05, 0.48)	0.27 (0.17, 0.43)	0.07 (0.02, 0.20)
		p=0.001	p<0.001	P<0.001

	Intended emotion			
Voices	Happy	Sad	Angry	Fearful
Controls	1	1	1	1
Early HD	0.53 (0.30, 0.95)	0.38 (0.19, 0.76)	0.34 (0.19, 0.62)	0.17 (0.09, 0.34)
	p=0.034	p=0.006	p<0.001	p<0.001

	Intended emotion			
Music	Happy	Sad	Angry	Fearful
Controls	1	1	1	1
Early HD	0.38 (0.17, 1.87)	0.35 (0.17, 0.73)	0.37 (0.22, 0.64)	0.45 (0.27, 0.75)
	p=0.021	p=0.005	p<0.001	p=0.002

An odds ratio <1 means that the early HD group is less likely to recognise stimuli correctly than controls; where 95% confidence intervals are both <1 this means that the odds ratio is statistically significant at the p<0.05 level.

APPENDIX 8: DETAILS OF REGRESSION MODELS USED IN PART 2

BACKGROUND DATA

TMT A, TMT B, TMT B-A and Peretz scores were all (separately) modelled as a function of group (Control or HD), controlling for the effects of age and IQ by including them as covariates (equation 1)

$$\text{Score} = \beta_1 \text{ group} + \beta_2 \text{ age} + \beta_3 \text{ IQ} + \mu + \varepsilon \quad (1)$$

where μ is a constant and ε is an error term.

Because of the non-normality of the data 95% accelerated bias-corrected bootstrap confidence intervals were estimated for the parameters for all three Trail-Making scores. Detailed results for each variable are shown in the following tables.

Regression coefficients for TMT A

TMT A	Observed coefficient (β)	95% accelerated bias-corrected bootstrap CI
Group	14.18	9.32, 21.44 ($p<0.05$)
Age	0.21	-0.01, 0.43 ($p>0.05$)
IQ	-0.06	-0.45, 0.17 ($p>0.05$)
Constant	19.50	-8.48, 64.77 ($p>0.05$)

Adjusted $R^2 = 0.29$

Regression coefficients for TMT B

TMT B	Observed coefficient (β)	95% accelerated bias-corrected bootstrap CI
Group	70.87	43.90, 108.82 ($p<0.05$)
Age	1.53	0.29, 2.73 ($p<0.05$)
IQ	-1.00	-2.47, 0.13 ($p>0.05$)
Constant	99.17	-32.76, 272.38 ($p>0.05$)

Adjusted $R^2 = 0.36$

Regression coefficients for TMT B-A

TMT B-A	Observed coefficient (β)	95% accelerated bias-corrected bootstrap CI
Group	56.69	29.50, 87.40 ($p<0.05$)
Age	1.32	0.15, 2.56 ($p<0.05$)
IQ	-0.94	-2.16, 0.04 ($p>0.05$)
Constant	79.67	-27.44, 205.03 ($p>0.05$)

Adjusted $R^2 = 0.32$

Regression coefficients for Peretz score

Peretz	Observed coefficient (β)	*t*(18)	95% CI
Group	-12.78	-2.85 (p=0.011)	-22.20, -3.37
Age	-0.11	-0.69 (p=0.50)	-0.43, 0.22
IQ	0.30	1.67 (p=0.11)	-0.08, 0.68
Constant	75.08	3.42 (p=0.003)	28.89, 121.27

Adjusted $R^2 = 0.38$

Age contributed significantly to TMT B and TMT B-A scores, and was only just non-significant for TMT A. IQ tended to have small, non-significant effects on outcome, although also was approaching statistical significance for TMT B-A. Nevertheless it was though important to keep both these predictors in all the models. Lack of evidence that a potential predictor affects outcome significantly can be due to lack of power, and even a predictor that does not reach statistical significance can materially affect the outcome variable in conjunction with other predictors. Johnson et al. (2007) make the point that variables such as age and IQ often have subtle effects on outcomes, but that most small studies are underpowered to detect them; they recommend that studies still include them as covariates because of this evidence. This also means that the effects reported here are comparable with those from other studies that also adjust for age and IQ.

EMOTION RECOGNITION DATA

Probability of a correct emotion recognition response was modelled as a function of group, modality, emotion and their interactions, controlling for the effects of age, IQ, Benton facial recognition ability, Trail-Making B-A score and gender by including them as covariates. Adjusted odds ratios for each emotion and modality are presented in the main body of the thesis. Conditional odds ratios for age, IQ, Benton, TMT B-A and gender are presented below.

Conditional odds ratios from the emotion recognition model

Emotion recognition	Odds ratio	95% CI
Age	0.98 (p=0.01)	0.96, 0.99
IQ	1.02 (p=0.016)	1.00, 1.03
Benton	1.03 (p=0.14)	0.99, 1.07
TMT B-A	1.00 (p=0.58)	0.996, 1.002
Gender	1.09 (p=0.57)	0.81, 1.47

In this model age and IQ had statistically significant effects on outcome. Benton, TMT B-A and gender did not. However theoretically it is known that gender and facial recognition ability have effects on facial emotion recognition (Henley et al., 2008; Johnson et al., 2007), and that executive function is likely to have an impact as well (Omar et al., 2011) and therefore, as above, it was thought appropriate to include these covariates.

APPENDIX 9: BATTERY ORDER

1. Trail-Making Test A & B
2. Benton Facial Recognition Test
3. Peretz Battery of Amusia
4. Multimodal Emotion Recognition
5. Hailstone Music Questionnaire
6. NART